ABSTRACT

Title of Dissertation: ALGORITHMIC DELEGATION:
FROM CLASSIC AUCTIONS
TO MODERN DIGITAL ECONOMY

SUHO SHIN
Doctor of Philosophy, 2026

Dissertation Directed by: Professor MohammadTaghi Hajiaghayi
Department of Computer Science
University of Maryland

Examining how a principal can effectively *delegate* tasks to an agent is a central problem in many real-world settings, and has been studied extensively in economics, computer science, and operations research. Representative examples include a team manager delegating problem-solving tasks to team members, and a funding agency selecting researchers to advance specific fields. More broadly, it subsumes a wide class of applications, ranging from traditional economic settings—where a public sector delegates trade governance to a self-interested broker—to the modern digital economy, where platforms such as YouTube rely on creators for content production.

In each case, the principal benefits from the agent's informational advantage, often lacking the expertise or time to make decisions independently. However, agents typically have objectives that are not perfectly aligned with those of the principal, giving rise to *moral hazard*, where the agent may act strategically for personal gain. The principal's challenge is therefore to design incentives that align the agent's behavior with their own objectives.

This article explores two fundamental questions in this context: (i) given a self-interested agent, how can the principal design an efficient mechanism that incentivizes the agent's behavior in favor of the principal, and what approximation guarantees can be achieved relative to the first-best (omniscient) outcome? and (ii) if the principal cannot intervene in the agent's decision-making, how much worse can the outcome be compared to the first-best outcome in hindsight?

I investigate these questions across four representative domains spanning both classical and modern applications: (i) delegation of choices in organizations, (ii) delegation of trade in auctions, (iii) delegation of production in modern online platforms, and (iv) delegation of information retrieval via generative AI. This work provides both theoretical insights and practical relevance within the broader literature on algorithmic game theory.

ALGORITHMIC DELEGATION:
FROM CLASSIC AUCTIONS TO MODERN DIGITAL ECONOMY

by

SUHO SHIN

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2026

Advisory Committee:
Professor MohammadTaghi Hajiaghayi, Chair/Advisor
Professor Lawrence Ausubel
Professor Ming Lin
Professor David Mount
Professor Negin Golrezaei
Professor Haifeng Xu

© Copyright by
Suho Shin
2026

To my parents, my brother, my grandmother, and my beloved puppy Moongchi.

Foreword

Delegation is a fundamental feature of modern decision-making process, where a principal relies on agents—humans, algorithms, or AI systems—to act on her behalf when she cannot observe all relevant information or actions in decision-making process. The principal delegates decision rights to agents who possess more information but may have misaligned objectives, preferences, or incomplete understanding of the principal's preferences. This misalignment between the principal's objective and the agent's incentives can lead to inefficiency compared to the *first-best* benchmark—namely, the outcome a fully informed principal would implement if she directly observed all signals.

A concrete instance arises in online content platforms such as YouTube, where the platform delegates content production to millions of independent creators. Each creator observes private information about their expertise and costs, and strategically adjusts behavior to gain exposure through the recommendation system. This is just one manifestation of a broader phenomenon: *delegation* arises across many domains. Managers depend on employees or algorithms to make operational choices, and individuals now delegate the task of information retrieval to large language models (LLMs). In each case, decisions are made indirectly—through intermediaries who act strategically.

While such delegation enables scalability and automation, it also introduces inefficiencies and ethical risks. Strategic or misaligned agents can exploit algorithmic systems, distort market outcomes, or propagate biased behavior. ***Understanding the limits and design principles*** of delegation mechanisms is therefore essential for building reliable platforms, markets, and human–AI ecosystems.

Despite decades of progress in mechanism and market design, we still lack an understanding of how to design decision systems that remain efficient when control is decentralized, information is incomplete, and agents behave strategically. This gap stems partly from the slow development of foundational theory in delegation up to seminal works [1–4], and partly from the emergence of new

applications—such as online content platforms and LLMs—that existing models cannot yet capture. **My research agenda investigates the theory of delegation from a modern algorithmic perspective**, offering insights into real-world decision-making and addressing theoretical gaps in the literature of algorithmic game theory, addressing two fundamental challenges in this context:

Question 1: **Efficient delegation mechanisms and approximation guarantees.** Given a self-interested agent, design a mechanism for the principal that optimally incentivizes the agent to act in the principal's interest, and quantify the approximation to the first-best (omniscient) outcome.

Question 2: **Fundamental loss without intervention.** When the principal cannot intervene in the agent's decision, analyze how much worse the outcome can be compared to the first-best outcome in hindsight.

My research investigates these questions in **four representative domains spanning classic and modern applications**: (1) *delegation of choices* in classic organizations, (ii) *delegation of trade* in auctions, (iii) *delegation of production* in modern online platforms, and (iv) *delegation of information acquisition* in the era of large language models (LLMs). These results characterize how much the principal's utility decreases compared to the first-best outcome and what efficient mechanisms can achieve near-optimal outcomes.

1. Delegation of Choices. A canonical model of delegation is the *delegated choice* problem [1, 2], where a principal delegates the selection of one among several uncertain options to an informed but self-interested agent. Each option i yields random utilities X_i for the principal and Y_i for the agent; the agent observes the realizations, but the principal only knows their distributions. The principal must thus design a *screening mechanism* that determines whether to accept the agent's proposal, for example by imposing a threshold on her own utility and accepting the agent's proposal when

it exceeds the threshold. This model captures many real-world situations—such as a manager approving a project proposed by an employee, a research foundation evaluating project proposals, or a hiring manager screening applicants.

[2] characterize the optimal mechanisms under stylized assumptions, and [3] show that the utility loss from delegation—the *delegation gap*—compared to the first-best is bounded by a constant factor using an elegant connection to the prophet inequality problem [5, 6]. However, these results are limited to *narrow settings* where the principal selects a single option and can freely inspect the agent's proposal. These assumptions are far from real-world applications, where decisions are often combinatorial or costly to verify. Thus, fundamental questions in this context are:

1. What if the principal needs to select a **set of options** that may not be a singleton, subject to a combinatorial set constraint that may govern the collection of feasible sets?

2. What if the principal needs to **inspect the agent's proposal at some cost**?

My recent papers address these questions as summarized in what follows.

In [7] (EC'25), I extend the delegated-choice framework to cases where the principal selects **multiple options** under a set constraint and show that the delegation loss remains small. First, I show that the connection to the prophet inequality persists if and only if the constraint is a matroid. Then, I study downward-closed constraints—a class that likely subsumes most practical constraints. Surprisingly, I devise a constant-factor approximate mechanism using a novel combination of set-wise thresholds and element-wise thresholds. Notably, this is the *first stark separation* between delegated choice and prophet inequality, as constant-factor approximation is shown to be impossible for prophet inequality [8].

In [9] (EC'25), I further analyze a setting where the principal **can inspect additional options**

at some cost. Notably, this setting *generalizes Pandora's box problem [10]* which has numerous applications in matching markets and search theory [11], establishing a new bridge between delegation, online decision-making, and search theory. I first prove that the optimal mechanism is NP-hard to compute, but also design a 3-approximation that combines a novel search algorithm with a delegation mechanism.

Many delegation mechanisms rely on prior knowledge of the utility distribution for each option. In practice, such information must be learned, prompting interest in *prior-independent mechanisms* that operate without explicit distributional knowledge. In [12] (AAAI'24), I study an online learning scenario where the principal repeatedly interacts with a single agent, using dynamic screening rules to elicit feedback and gradually estimate the utility distribution. I propose efficient algorithms with sublinear regret for both myopic and farsighted agents, who may optimize either short-term or long-term utility. In contrast, in [13] (EC'23), I show that prior learning may be unnecessary if the principal can delegate to multiple agents. When agents derive utility only if their option is selected, competitive tension emerges, motivating agents to reveal stronger proposals. This mechanism achieves improved outcomes without prior knowledge, though its effectiveness crucially depends on agents' information and heterogeneity.

2. Delegation of Trade. Auctions are perhaps the oldest market structure to exchange items. A classic setting in auctions is the bilateral trade problem, where a strategic buyer and seller aim to trade a single item. The seminal result by [14] shows that it is impossible to achieve ex-post efficiency under the standard constraints of (i) incentive compatibility, (ii) individual rationality, and (iii) budget balance. This impossibility has led to a long line of research seeking approximate efficiency, particularly in terms of gains from trade (GFT)—the net increase in social welfare from the

trade. A major open question was whether a constant-factor approximation to the first-best GFT is achievable, until resolved affirmatively by [15].

On the other hand, in many real-world scenarios—such as financial markets, real estate, over-the-counter markets, and markets for demerit goods with licensed intermediaries—the governance of trade is delegated to a **self-interested broker** whose objective is not to maximize social welfare. Despite its practical relevance and strong theoretical foundation, there has been surprisingly little work on bilateral trade with a broker. Accordingly, it remains questionable whether the presence of a self-interested broker interferes with the possibility of achieving a constant-factor approximation to the first-best GFT. Thus, I have been investigating the following research questions:

1. Does the presence of a **self-interested broker** interfere with the possibility of achieving constant-factor approximation to the first-best GFT? What happens when the broker **has only limited data samples** about buyers' and sellers' distributions?

2. Can we obtain a **Bulow–Klemperer style result** [16] showing that recruiting a small number of additional participants can offset the broker's informational disadvantage?

In [17] (SODA'25), I initiate the study of bilateral trade with a strategic broker. I prove that if the broker runs a profit-maximizing mechanism, **approximation to the first-best GFT becomes impossible**, in contrast to the benevolent-broker case of [15]. However, when the broker is restricted to run posted-pricing mechanisms—prevalent in real markets due to simplicity—I show that a constant-factor approximation is achievable under standard regularity assumptions.

In [18] (WWW'26), I study the *sample-limited* setting where the broker has access to a few independent samples from the buyer's and seller's distributions. I show that, unlike the fully informed case, **a single-sample mechanism** can already achieve constant-factor approximations for both the

broker's profit and the overall GFT, providing a simple and prior-independent design. Interestingly, I observe that the approximation factor for the GFT remains almost the same as in the scenario without a broker [19]. From a practical perspective, a policymaker can *restrict the intermediary's information to improve social welfare.*

3. Delegation of Production. In the modern digital economy, online platforms such as YouTube delegate the role of producing content to individuals. Producing high-quality content, however, requires significant time and effort. To keep content producers engaged in the ecosystem, platforms often provide monetary compensation or encourage them to embed advertisements in their content. Crucially, the incentives for content producers are closely tied to the recommendation algorithm itself, as greater visibility to users translates to higher potential earnings.

Despite the strategic nature of content production, most recommendation algorithms *do not account for the producers' strategic behavior*—namely, their efforts to maximize revenue while managing production costs. This interaction is critical to designing an efficient ecosystem, yet it remains underexplored in existing models. Thus, the fundamental questions in this context are:

1. How can a platform optimally incentivize content producers when they strategically choose the **quality** of their content to maximize expected rewards minus production costs?

2. What if producers also **strategize over quantity**, deciding how many pieces of content to produce?

In [20] (STOC'26), I study a *contest design problem* where a platform operates a recommender system with n strategic producers who select **costly effort levels to generate content of a given quality**. The platform recommends the i-th highest-quality content with probability p_i, aiming to balance user welfare and content quality at equilibrium. This induces an optimization problem over

a convex-plus-concave objective that is typically intractable. Interestingly, I prove that the optimal policy always satisfies $p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$, implying that *fairness in exposure*, often imposed as an external constraint, emerges endogenously. Using this structure, I show that the optimal policy can efficiently be computed.

En route to these results, I observe a *surprising reduction* from a structured high-dimensional nonconvex optimization to a single-dimensional optimization by connecting the shape of the gradient sequences of the objective function to the number of transition points in optimum for a constrained symmetric function optimization problem, which may be of broader interest to the optimization community. This can be also viewed as a *generalization of the Schur-convexity* [21], a central concept in numerous optimization problems.

In [22,23] (AISTATS'22, AAAI'25 oral), I study an extreme version of **strategic behavior over quantity**. Here, producers can replicate their content freely to exploit exploration phases of recommendation algorithms. I design algorithms that deter such replication while maintaining near-optimal regret in the multi-armed bandit framework,[1] while showing that standard algorithms are vulnerable to careless replication of content. A core idea is to fairly divide the exploration amount between producers. In other words, fairness again implies robustness against such manipulation.

4. Delegation of Information Acquisition. In recent years, generative AI—especially large language models (LLMs)—has transformed how people gather information, with many now using LLMs instead of search engines [24]. Yet LLMs often misinterpret user intent or generate misaligned outputs due to ambiguous prompts or limited awareness of user preferences. Thus, a key

[1] Interestingly, the motivation behind this question comes from my industry experience. When I was working as a data scientist in an advertisement platform company, I observed an anomaly that small-business advertisers register the same ads multiple times. After thorough data analysis, I observe that they are abusing exploration resources of recommendation algorithms. Notably, this even motivates my journey on Ph.D. to study incentives in markets.

challenge for improving user experience is how to effectively control LLM outputs to align with individual preferences.

A standard approach is reinforcement learning from human feedback (RLHF), where LLMs are post-trained using large datasets of outputs and human evaluations. However, this process is usually costly and does not scale to provide personalized alignment, as retraining per user is too costly. Recently, there has been growing attention to *decoding-time alignment*,[2] which aims to align outputs during inference time without retraining. However, this emerging direction lacks theoretical foundations and often relies on ad hoc external modules. In this context, I have been investigating the following set of questions:

1. How can we align an LLM's output *during decoding* through **repeated interactions and personalized feedback**, without retraining?

2. In practice, LLMs rely on retrieval-augmented generation (RAG); how can we **optimally learn to retrieve** document sets to improve accuracy?

3. How should platforms **incentivize and finance** high-quality document provision while covering operational costs?

In [25] (ICML'25), I study *decoding-time alignment*: aligning outputs during inference time from **repeated user feedback**. I introduce the *tokenized bandit* framework, extending multi-armed and linear contextual bandits to sequential token generation where feedback arrives only after the full sequence. I establish exponential lower bounds and, under a natural structural condition, design algorithms with sublinear regret. As a byproduct, I show that *greedy decoding is provably near-optimal*, offering the *first theoretical justification* for a widely used decoding heuristic.

[2]Decoding (inference) algorithms decide which token to select given a probability distribution over tokens from the transformer.

In my position paper [24] (SIGecom Exchanges 22.2), I argue that LLMs can embed textual advertisements within outputs to **finance operational costs**, contrasting this with traditional online ads. In [26] (NeurIPS'24), I develop an auction framework for RAG where retrieved documents are allocated proportionally to bids; I show how this rule can be implemented via discrete-choice methods and that it optimizes a *logarithmic social welfare* objective that balances allocative efficiency and fairness among advertisers. In [27] (ICLR'26 Workshop on AI for Mechanism Design and Strategic Decision Making), I study a two-sided market in which the platform **earns revenue from advertisers** and **compensates document providers** under a budget-balance constraint, and analyze mechanisms that maximize user utility. Notably, this *generalizes the standard double auction beyond linear utility objectives*.

Acknowledgments

First and foremost, I offer my deepest gratitude to my advisor, MohammadTaghi Hajiaghayi, whose intellectual generosity, discerning guidance, and unwavering mentorship have shaped every stage of my research journey. His steadfast support has been indispensable to my growth—not only as a scholar, but as a thinker. I am grateful for his support in all aspects of my life, even beyond academia. This thesis would not have been possible without his guidance, and I will always be deeply indebted to him.

I cannot fully express how grateful I am to Negin Golrezaei for her support and mentorship. Her generosity, thoughtfulness, and inspiring research guidance have had a profound impact on me. I am deeply grateful to her for hosting me at MIT and for the invaluable career advice she has shared. I am also deeply thankful to Sébastien Lahaie for his mentorship, warm support, and the exciting opportunities I had at Google Research. I am greatly thankful to Haifeng Xu for his mentorship, thoughtful career advice, and for hosting me at the University of Chicago. I also thank Piotr Krysta for his engaging research discussions, thoughtful mentorship, and for hosting me on several occasions over the years.

I am profoundly grateful to my dissertation committee—Lawrence Ausubel, Negin Golrezaei, MohammadTaghi Hajiaghayi, Ming Lin, David Mount, and Haifeng Xu—for their time, thoughtful critiques, and generous counsel. I am additionally thankful to the wonderful staff at UMD—especially Tom Hurst, Migo Gui, Jodie Gray, and Sharron McElroy—for their support and for the exceptional work they do.

I owe sincere thanks to my collaborators, colleagues, and coauthors, whose intellectual companionship and shared curiosity have enriched this work in ways beyond enumeration: Kiarash Banihashem, MohammadHossein Bateni, Simina Branzei, Elette Boyle, Shayan Chashm Jahan, Hoyong

Choi, Vincent Cohen-Addad, Seyed Esmaeili, Soheil Feizi, Dimitris Fotakis, Ilya Hajiaghayi, Stanisław Kaźmierowski, Joosung Lee, Mohammad Mahdavi, Edgar Martinez, Jungseul Ok, Gary Peng, Reed Phillips, Keivan Rezaei, Mohammad Sharifi, Alex Slivkins, Max Springer, Amos Stern, Yung Yi, Chenghao Yang, and Kun Wang. Their insights and encouragement have left an enduring imprint on this dissertation.

I extend particular thanks to my dear labmates Mohammad Mahdavi and Keivan Rezaei for their encouragement, kindness, and remarkable generosity with their time—both within and beyond academic life—especially during difficult periods shaped by uncertainty and geopolitical hardship. I am likewise grateful to Kiarash Banihashem and Max Springer for their thoughtful guidance, which proved invaluable throughout my graduate studies.

My heartfelt thanks go to my dear friends in South Korea: Daewon Kim and Minsik Kim for their warm support and encouragement, and ByoungUk Lee and Jeonghyun Song for always making time for me—for the laughter and moments of joy we shared despite the distance. I also thank my dear UMD friends Joohyung Song and Sungsu Kwak, whose companionship brought lightness and welcome moments of ease during even the most demanding periods of my Ph.D. journey. I am deeply appreciative of my roommate, Keonwoo Oh, for his constant consideration and generosity in our shared daily life. I thank my UMD classmates Daeun Jung, Nakyung Lee, and SeongJong Yoo for their steady emotional and logistical support throughout my graduate studies. I am grateful as well to my tennis mates—Jiwoong Jang, Yeleen Jo, Brian Kim, Euna Lee, Kihoon Yuk, and all the members of the UMD Korean Graduate Tennis Club—for the many spirited moments and the lasting memories.

Last but not least, no words can adequately convey the depth of my gratitude to my family—my parents, my brother, my grandmother, and my beloved puppy Moongchi—all of whom have been both my steadfast foundation and my greatest teachers in life. Special congratulations to my dear brother as

he begins a new chapter of his life in marriage; may it be filled with lasting happiness and blessings. Every step of my Ph.D. journey rests upon the love and sacrifice of my family. *This thesis is, in spirit, theirs.*

Table of Contents

Dedication ii

Foreword iii

Acknowledgements xii

Table of Contents xv

List of Tables xxi

List of Figures xxii

List of Algorithms xxiii

I Delegation of Choices 1

Chapter 1: Delegated Choice with Combinatorial Constraints 4
- 1.1 Introduction . . . 4
- 1.2 Our results and techniques . . . 7
  - 1.2.1 Agent-oblivious mechanisms for downward-closed constraints . . . 8
  - 1.2.2 Agent-oblivious mechanisms for matroid constraints . . . 12
  - 1.2.3 Agent-aware mechanisms . . . 16
- 1.3 Model . . . 18
- 1.4 Preliminaries . . . 21
  - 1.4.1 Matroids . . . 21
  - 1.4.2 Prophet inequality . . . 22
  - 1.4.3 Implementation . . . 25
- 1.5 Downward-closed constraint . . . 27
- 1.6 Matroid constraint . . . 30
  - 1.6.1 Implementation from matroid prophet inequality . . . 31
- 1.7 Necessity of matroid constraint for implementability . . . 32
- 1.8 Implementation of broader class of non-adaptive algorithms . . . 33
- 1.9 Agent-aware mechanisms . . . 35
  - 1.9.1 Power of knowing the arrival order is negligible . . . 36

Chapter 2: Delegation with Costly Inspection 37
- 2.1 Introduction . . . 37
- 2.2 Our results and techniques . . . 39
  - 2.2.1 First attempt: PNOI policy or SPMI . . . 41
  - 2.2.2 Next attempt: using ideas of [28, 29] . . . 43
  - 2.2.3 Costless delegation . . . 45

2.2.4 Costly delegation . . . 48
2.2.5 Related Works . . . 49
2.3 Problem Setup . . . 50
2.3.1 Delegated Choice . . . 50
2.3.2 Delegated Choice Problem with Inspection Cost . . . 52
2.4 Main Results . . . 57
2.4.1 Warm-up . . . 57
2.4.2 3-approximation in costless delegation setting . . . 58
2.4.3 Constant factor approximation for general setting . . . 68
2.5 Remaining Proofs . . . 71
2.5.1 Proof of Proposition 2.3.6 . . . 71
2.5.2 Proof of Proposition 2.4.1 . . . 72
2.5.3 Proof of Proposition 2.4.2 . . . 74
2.5.4 Proof of Proposition 2.4.3 . . . 74
2.5.5 Proof of Proposition 2.4.10 . . . 75

Chapter 3: Delegating to Multiple Agents 75
3.1 Introduction . . . 76
3.2 Overview of results . . . 80
3.2.1 Contributions and techniques . . . 81
3.3 Related work . . . 88
3.4 Problem setup . . . 90
3.4.1 Further preliminaries . . . 93
3.5 Complete information . . . 96
3.5.1 Bayesian mechanism . . . 97
3.5.2 Prior-independent mechanism . . . 99
3.6 Incomplete information . . . 104
3.7 Examination cost . . . 109

II Delegation of Production 112

Chapter 4: Optimal Contest for Online Platforms 113
4.1 Introduction . . . 114
4.1.1 Results and Techniques . . . 118
4.2 Related Works . . . 130
4.3 Model . . . 131
4.4 Existence and Uniqueness of Symmetric MNE . . . 136
4.4.1 Existence . . . 136
4.4.2 Uniqueness . . . 139
4.5 Detaching Equilibrium Dependency in Optimization . . . 140
4.5.1 HARDMAX can be Arbitrarily Bad . . . 142
4.6 Optimal Policy Structure for Single-Minded RS . . . 142
4.7 Optimal Policy Structure for Convex-minded RS . . . 144
4.7.1 Proof Sketch of Theorem 4.1.3 . . . 145

4.7.2 Implication on High-dimensional Nonconvex Optimization Problem 150
4.8 Computing Optimal Policy 152
4.9 Beyond RS: General Objective Functions 156

Chapter 5: Replication-proof Bandit Mechanism Design with Bayesian Agents 160
5.1 Introduction 160
5.1.1 Contributions and Techniques 163
5.2 Model 165
5.3 Failure of Existing Algorithms 170
5.4 Warm-up: Single-agent Setting 171
5.5 Multi-agent Replication-proof Algorithm 178

III Delegation of Trade 181

Chapter 6: Gains-from-trade in Bilatreral Trade with a Broker 183
6.1 Introduction 183
6.1.1 Our Results 184
6.1.2 Further Related Work 193
6.2 Problem Setup 195
6.3 Posted-Pricing Mechanisms 200
6.4 Optimal Mechanisms 210

Chapter 7: Single-sample Bilateral Trade with a Broker 225
7.1 Introduction 225
7.1.1 Our Results and Techniques 228
7.1.2 Related Work 235
7.2 Model 236
7.3 Preliminary Lemmas 240
7.4 Symmetric Case 242
7.4.1 Lower Bounds 244
7.4.2 Upper Bounds 247
7.5 Stochastic Dominance Case 252
7.5.1 Lower Bounds 252
7.5.2 Upper Bounds 259

IV Delegation of Information Acquisition 261

Chapter 8: Tokenized Bandit for LLM Decoding and Alignment 263
8.1 Introduction 263
8.1.1 Overview of results, assumptions, and implications 264
8.2 Problem Setup 267
8.3 Tokenized Linear Bandit 270
8.3.1 Diminishing distance with more commons (DDMC) 271
8.3.2 Excessive Optimism under the Face of Uncertainty 273

8.4 Tokenized Multi-armed Bandit . . . 276
8.5 Applications and Experiments . . . 278
8.5.1 Application: LLM alignment . . . 278
8.5.2 Implication: LLM decoding . . . 279
8.5.3 Validating DDMC . . . 281

Appendix A: Delegated Choice with Combinatorial Constraints 1
A.1 Further Preliminaries on Matroids . . . 1
A.2 Implementation of Agent-aware Mechanisms . . . 2
A.3 Some Probability Tools and Inequalities . . . 3
A.4 Proofs . . . 8
A.4.1 Proof of Theorem 1.5.1 . . . 8
A.4.2 Proof of Theorem 1.6.1 . . . 14
A.4.3 Proof of Corollary 1.6.3 . . . 17
A.4.4 Proof of Theorem 1.7.2 . . . 18
A.4.5 Proof of Proposition 1.8.1 and Corollary 1.8.2 and 1.8.3 . . . 22
A.4.6 Proof of Theorem 1.8.4 . . . 25
A.4.7 Proof of Proposition A.2.1 and Corollary A.2.2 and A.2.3 . . . 30
A.4.8 Proof of Theorem 1.9.1 . . . 31
A.4.9 Proof of Proposition 1.4.2 . . . 37

Appendix B: Delegating to Multiple Agents 39
B.1 Numerical Experiments on the Tightness of Theorem 3.6.7 . . . 39
B.2 Further discussion on feasible mechanisms . . . 39
B.3 Bayes Correlated equilibrium . . . 40
B.4 Bulow-Klemperer-style result for prior-independent mechanism . . . 42
B.5 Comparing multi-agent and single-agent delegation . . . 43
B.6 Proofs . . . 46
B.6.1 Proof of Theorem 3.4.11 . . . 46
B.6.2 Proof of Theorem 3.5.1 . . . 47
B.6.3 Proof of Theorem 3.5.3 . . . 50
B.6.4 Proof of Theorem 3.5.5 . . . 52
B.6.5 Proof of Theorem 3.5.8 . . . 54
B.6.6 Proof of Theorem 3.5.10 . . . 55
B.6.7 Proof of Theorem 3.5.11 . . . 56
B.6.8 Proof of Lemma 3.5.12 . . . 57
B.6.9 Proof of Lemma 3.5.14 . . . 60
B.6.10 Proof of Theorem 3.5.15 . . . 61
B.6.11 Proof of Lemma 3.5.16 . . . 62
B.6.12 Proof of Lemma 3.5.17 . . . 65
B.6.13 Proof of Theorem 3.5.18 . . . 66
B.6.14 Proof of Theorem 3.6.4 . . . 67
B.6.15 Proof of Theorem 3.6.7 . . . 69
B.6.16 Proof of Theorem 3.7.5 . . . 78
B.6.17 Proof of Theorem B.3.2 . . . 82

B.6.18 Proof of Theorem B.5.3 . . . 87

Appendix C: Replication-proof Bandit Mechanism Design with Bayesian Agents 89
C.1 Related Works . . . 89
C.2 Motivations . . . 89
C.3 Extension to the Asymmetric Setting . . . 90
C.4 Failure of Existing Algorithms . . . 91
C.5 Proof of Proposition 5.4.5 and Theorem 5.4.9 . . . 102
C.6 Proof Sketch of Proposition 5.4.10 . . . 107
C.7 Proof of Theorem 5.5.1 . . . 108
C.8 Proof of Theorem 5.5.2 . . . 118

Appendix D: Optimal Contest beyond Convexity 122
D.1 Preliminaries for Section 4.4 . . . 122
D.1.1 Preliminaries on Topology . . . 122
D.1.2 Preliminaries on Games and Existential Results . . . 128
D.2 Omitted Proofs in Section 4.4 . . . 131
D.2.1 Proof of Theorem 4.4.2 . . . 131
D.2.2 Proof of Proposition 4.4.1 . . . 137
D.2.3 Proof of Proposition 4.4.3 . . . 137
D.2.4 Proof of Lemma 4.4.4 . . . 139
D.2.5 Proof of Theorem 4.4.5 . . . 140
D.3 Omitted Proofs in Section 4.5 . . . 141
D.3.1 Proof of Theorem 4.1.1 . . . 141
D.3.2 Proof of Proposition 4.5.1 . . . 153
D.4 Omitted Proofs in Section 4.6 . . . 155
D.4.1 Proof of Lemma 4.6.5 . . . 155
D.4.2 Proof of Theorem 4.1.2 . . . 159
D.5 Omitted Proofs in Section 4.7 . . . 160
D.5.1 Proof of Lemma 4.7.2 . . . 160
D.5.2 Proof of Lemma 4.7.3 . . . 166
D.5.3 Proof of Theorem 4.1.3 . . . 167
D.6 Omitted Proofs in Section 4.8 . . . 173
D.6.1 Proof of Lemma 4.8.1 . . . 173
D.6.2 Proof of Lemma 4.8.2 . . . 173
D.6.3 Proof of Lemma 4.8.4 . . . 175
D.6.4 Proof of Theorem 4.8.6 . . . 175
D.6.5 Proof of Lemma 4.8.7 . . . 176
D.7 Omitted Proofs in Section 4.9 . . . 176
D.7.1 Proof of Theorem 4.1.5 . . . 176
D.7.2 Proof of Corollary 4.9.1 . . . 179
D.7.3 Proof of Corollary 4.9.3 . . . 181

Appendix E: Gains-from-trade in Bilateral Trade with a Broker 183
E.1 Proof of Theorem 6.1.8 . . . 183

E.1.1 Lower bounds . . . 183
E.1.2 Derivations for Table E.2 and E.3 . . . 198
E.1.3 Upper Bounds . . . 202
E.2 Proof of Theorem 6.4.6; Public-Buyer . . . 203

Appendix F: Single-Sample Bilateral Trade with a Broker 207
F.1 Remaining Proofs for Preliminary Lemmas . . . 207
F.1.1 Proof of Lemma 7.3.2 . . . 207
F.1.2 Proof of Lemma 7.5.1 . . . 211
F.2 Preliminaries for Upper-Bound Proofs . . . 212
F.3 Remaining Proofs for Symmetric Agents . . . 224
F.3.1 Proof of Theorem 7.4.3 . . . 224
F.4 Remaining Proofs for Asymmetric Agents . . . 225
F.4.1 Proof of Theorem 7.5.5 . . . 225
F.4.2 Proof of Theorem 7.5.6 . . . 226

Appendix G: Tokenized Bandit for LLM Decoding and Alignment 229
G.1 Related work . . . 229
G.2 k-DDMC and Lookahead Decoding . . . 230
G.3 Connection to Bandit Tree Search problem . . . 232
G.4 Proofs . . . 236
G.4.1 Proof of Proposition 8.3.3 . . . 236
G.4.2 Proof of Theorem 8.3.9 . . . 237
G.4.3 Proof of Theorem 8.5.1 . . . 248
G.4.4 Proof of Proposition 8.4.1 . . . 249
G.4.5 Proof of Theorem 8.4.2 . . . 249

Bibliography 253

List of Tables

3.1 Summary of results on Bayesian mechanism. *Symmetric* refers to the symmetric agents and α-*decaying* denotes the distributions such that its c.d.f. F satisfies $F(x) \leq x^\alpha$ for some $\alpha > 0$. Lower bounds for the symmetric setting remain as open problems. 83

3.2 Summary of results on prior-independent mechanism. *MHR* refers to MHR distribution, *Inc* refers to nondecreasing p.d.f., and *Sym Ind* refers to independent utility setting with the uniform distribution $U[0, 1]$. *Open* denotes that this regime remains an open problem. For general case lower bound for both settings, and symmetric case upper bound in the complete information setting, we assume that the support of the principal's utility function lies in $[0, L]$. For Sym MHR, Sym Inc, and lower bound in the incomplete information setting under Sym, we present the results for $U[0, 1]$ for brevity. 86

4.1 Optimal policy under different values of α (objective) and β (cost curvature). "Any" in the optimal policy column denotes that any policy with $p_n = 0$ would induce the same objective values. 120

7.1 Our results for the symmetric setting. Note that the approximability of the optimal profit is not applicable to the setting without a broker (or equivalently, with a benevolent broker). However, an upper bound on the approximation factor to the optimal profit remains an open problem in our setting with a broker. 230

7.2 Our results for the setting where F stochastically dominates G. An upper bound on the approximability of the first-best social-welfare in the stochastic dominance setting without a broker is unresolved to the best of our knowledge, although [30] show an upper bound of $1/2$ for general distributions in the setting without a broker. However, since [30] use an optimization problem to obtain the exact worst-case approximation ratio, it is unclear whether there exists a problem instance where F stochastically dominates G that achieves the $1/2$ bound. 233

A.1 Competitive ratio for small k in Corollary 1.6.3 and Lemma A.4.2, obtained by [31]. . 17

E.1 Summary of Lower Bounds 184

E.2 First-Best Gains-from-Trade by Case 186

E.3 Gains-from-Trade of Broker's Optimal Mechanism by Subcase 187

E.4 Examples of Values of a and b that Match Lower Bounds 202

List of Figures

4.1 Structures of optimal policy with respect to different α and β values given $n = 5$ producers. We compute the optimal policy using a brute-force search over every possible policy after discretizing the policy space with granularity 0.005, and using trapezoidal rule to approximate integral values over the grid $[0, 0.001, 0.002, \ldots, 1]$. For the left figure, we set $\beta = 2$ and for the right figure, we set $\alpha = 0.24$. This yields that our characterization is tight in a sense that the structures $p_1 > p_2 = \ldots = p_{n-1} \geq p_n = 0$ or $p_1 = p_2 = \ldots = p_{n-1} \geq p_n = 0$ indeed appear in the optimal policies. 122

4.2 Structures of optimal policy with respect to different α and β values given $n = 5$ producers. Darker red means larger p_1 values, and darker blue means larger $p_2 = p_3 = \ldots = p_{n-1}$ values. Each of the optimal policy is computed in a brute force manner after discretizing policy space with granularity 0.001 and using trapezoidal rule with number of points 200. Parameters α and β are selected from the ranges $[0.05, 1]$ and $[0.1, 5]$, respectively, with 1000 points calculated for each. 125

4.3 Behavior of d_i for $n = 5$. Since $p_5 = 0$, d_5 is omitted. Quasiconvexity of d_i follows from the quasiconvexity of $q(x)$ shown in Figure 4.4. 147

4.4 Behavior of $q(x)$ for $n = 5$ (see Appendix D.5.3 for definition). The quasiconvexity of $q(x)$ transfers to d_i due to variation diminishing properties of the kernel defined by $a_i(x)$. 150

8.1 The first two show distance gap under TruthfulQA and the next two is under HH-RLHF, with distance function d_1 and d_2. 279

B.1 Probability of the event E defined in (B.9). Lowerbound refers to the lower bound of the event E that we obtained in (B.15), and n and k denotes the number of agents and the number of solutions per agent. We note that x-axis is presented in the base of the natural number e. 40

C.1 Bad instance for UCB in Theorem C.4.1. Each node denotes the sequence of the bandit algorithm's selection for each instance. Blue node denotes the original arm with mean reward 1, and green one is the replica of it. Red nodes the original arm with mean reward 0, and yellow node is the replica of it. Note that at round $t = s_1 + s_2 + 2$, instance A realizes reward 0 two times, whereas B and C realizes three times, thus having 1.5 times in average. 93

G.1 Illustration of reductions. The above figure is an example of reduction from BTS to TMAB, and the below is from TMAB to generalized BTS (with k-ary tree). In both figures, blue nodes represent leaf nodes and orange node denotes `EOS` token. Dotted edge in the rightmost graph implies that the sequence contains such edge would have utility zero. Thus, only the sequences without dotted lines have (potentially) nonzero utility corresponding to the utility at the leaf node in the original instance of BTS. In the below figure, each path to `EOS` in the instance of TMAB becomes a leaf node in the instance of generalized BTS. 234

List of Algorithms

1 Active-Set Branch-and-Bound . . . 155

2 EOFUL . . . 274
3 GreedyETC . . . 277

4 Greedy . . . 1

5 UCB [32] . . . 91
6 H-UCB [22] . . . 91
7 ETC . . . 102
8 Prior-indepdenent ETC . . . 107
9 Hierarchical ETC with Restarting (H-ETC-R) . . . 108

10 k-Lookahead-EOFUL . . . 231
11 k-LookaheadETC . . . 231

Part I

Delegation of Choices

In this part of the article, results on several extensions of the fundamental problem called *delegated choice* are presented. Each chapter consists of a self-contained paper.

In the realm of public decision-making, a decision-maker frequently faces a task that demands expertise beyond its own, which necessitates the delegation of the decision-making process to an agent with expertise. Consider, for example, a scenario where a mayor (principal) in a local government wishes to undertake infrastructural development by paving roads that span a given set of cities. The mayor aims to minimize the total cost of paving the roads. However, she cannot directly estimate the potential cost of each edge in the graph, but only knows a distribution from which the cost of paving each road might be sampled, possibly inferred from historical data. Therefore, the mayor delegates the decision-making to an expert (agent) who can observe the exact cost of the edges. The expert, however, has a misaligned preference for paving roads that benefit their personal connections or interests. The mayor should now commit to a screening mechanism that filters out bad proposals from the expert. Then, a natural question is to design an efficient mechanism to maximize the principal's expected payoff against the selfish agent.

The *delegated choice* problem, dating back to [1,2], received a lot of attention from both the computer science and economics community, being one of the most fundamental principal-agent problems. In delegated choice, there is a set of pairs of mutually independent random variables $\{(X_i, Y_i)\}_{i\in[n]}$, where X_i and Y_i denotes a principal's and an agent's utility for element i, respectively. The principal wants to select an element that maximizes her[3] own utility X_i, however, she cannot access the value of the random variables directly. Instead, she delegates the role of examining the random variables to the agent, who proposes one of them to the principal. The agent observes the realization of the random variables at no cost, but has ulterior motives to maximize his own utility Y_i. To mitigate the agent be-

[3]Throughout, feminine pronouns refer to the principal and masculine ones refer to the agent.

ing too selfish, the principal commits to a screening mechanism that screens out the proposed element based on its utility. If the principal knows the distribution from which the utility of the solutions are drawn, the principal faces the problem of designing an optimal Bayesian mechanism to compete with the expected utility from the optimal solution in hindsight, *i.e.,* $\mathbb{E}\left[\max_{i\in[n]} X_i\right]$.

On the other hand, initiated by [5], the *prophet inequality* problem has attracted significant interest in the past few decades, with its fruitful connections to various problems spanning online auctions for digital goods [33], multi-dimensional mechanism design [34], stochastic optimization [35], secretary problem [36, 37], and online Pandora's box [38]. In the prophet inequality problem, n mutually independent non-negative random variables $W_1, W_2, \ldots, W_n$ with known distributions arrive one at each round, and the decision maker needs to irrevocably choose one random variable to compete against the maximal value in hindsight, *i.e.,* $\mathbb{E}\left[\max_{i\in[n]} W_i\right]$. The surprising result by [5] proved that there exists a stopping rule τ such that

$$\mathbb{E}\left[W_\tau\right] \geq \frac{1}{2}\mathbb{E}\left[\max_{i\in[n]} W_i\right],$$

which is known as a prophet inequality thus far. [6] flourished their result by showing that a simple threshold-based rule achieves the inequality above, and a numerous number of papers extend our understanding of the prophet inequality problem in various settings, *e.g.,* [39–42], or see [43] for a wonderful survey. Perhaps two of the most groundbreaking results are the *matroid* prophet inequality by [44] and *downward-closed* prophet inequality by [8].[4]

[4]To provide more context, [44] provided an optimal adaptive algorithm with $1/2$-approximation if the decision maker can select multiple elements subject to a matroid constraint, and the benchmark is the maximal independent set with maximum weight in hindsight. This, in particular, is surprising given that the matroid analogue of the secretary problem which has an intimate connection to the prophet inequality, *i.e.,* matroid secretary by [45], has not yet resolved over the decades. [8] obtained a $1/\log n$-competitive algorithm and $\log\log n/\log n$ lower bound for downward-closed prophet inequality, which is the first known result beyond the matroid constraint.

Quite recently, [3] observed a surprising connection between the prophet inequality and the delegated choice problem. They showed that the delegated choice problem is *equivalent* to the prophet inequality problem with oblivious stopping rules,[5] thereby constructing several efficient delegation mechanisms in various settings, especially a $1/2$-approximate Bayesian mechanism which is proven to be optimal. This connection, correspondingly, later extended to variants of the delegation problem such as delegated stochastic probing [46] and delegated pandora's box [47], all of which reduce the delegation mechanism design to prophet inequality.

Chapter 1: Delegated Choice with Combinatorial Constraints

This chapter presents the results of [7] (EC 2025).

1.1 Introduction

A natural follow-up question of [3] is, to what extent such connection between delegated choice and prophet inequality carries over, *i.e.,*

Does the connection between prophet inequality and delegated choice carry over to arbitrary combinatorial constraints?

Surprisingly, we show that this is not true for downward-closed constraints but holds if and only if the underlying constraint is matroid, revealing its first intrinsic *distinction* between the two problems

[5] An oblivious stopping rule sets a non-adaptive threshold for each element, and selects the element if and only if the element's value is at least that threshold, and is feasible with respect to the set constraint upon the currently selected set.

under combinatorial constraints.

Formally, we study *delegated choice with combinatorial constraints*, generalizing the delegated choice problem into multiple choices with set constraints.[1] More specifically, given a universe U of n elements, there is an underlying set system $\mathcal{M}$ over U which determines a collection of feasible sets $\mathcal{I} \subseteq 2^U$. Both the agent and the principal realize non-negative utility only if a feasible set is selected as a final outcome, and their utility is additive over the selected elements. The principal aims to design a Bayesian mechanism that competes with the optimal feasible set in hindsight, *i.e.*, $\mathbb{E}\left[\max_{I\in\mathcal{I}} \sum_{e\in I} X_e\right]$. We focus on two classes of agent-aware/oblivious mechanisms, depending on whether the principal knows the agent's utility distribution or not.

To highlight some of the main implications of our results:

1. We first prove that the reduction presented in [3] no more works for general downward-closed constraints. Nevertheless, we obtain a constant-approximate mechanism for downward-closed constraints, which implies a *strict separation* between prophet inequality and delegated choice since prophet inequality does not admit constant approximation due to [8]. The mechanism here exploits both the bundle-wise thresholds and element-wise thresholds inspired by multi-dimensional auction design, which has not been studied in the context of delegated choice and its variants.

2. For matroid constraints, however, we show that the connection between two problems carries over *up to matroid*, by establishing the connection between the agent's best-response and the greedy algorithm's optimality in the matroid structure. We further prove that the matroid constraint is a *necessary* condition for such correspondence, which complements our result on the

[1] This problem is also studied as a special case of delegated stochastic probing by [46], but from a different perspective.

separation under downward-closed constraints.

3. Finally, we examine the power of agent-aware mechanisms that know the agent's utility distribution. We obtain a natural connection to prophet inequality which knows the distribution of the arrival order, defined as an order-hinted algorithm. Interestingly, however, we show a minimax-style result, implying that order-hinted non-adaptive algorithms (and induced mechanism) do *not have strictly better* competitive ratios compared to order-oblivious non-adaptive algorithms (and induced mechanism), which might be of independent interest.

Further, it is worth noting that the strict separation between delegated choice and prophet inequality for downward-closed constraints is *not expected* at first glance. Indeed, each problem has its own advantages in designing algorithms/mechanisms when compared with each other.[2] In the prophet inequality, one can design an *adaptive algorithm* that determines its decision at each round based on what it has observed before and its corresponding previous decision. In the delegated choice, one cannot design a mechanism that adaptively changes its behavior on the remaining elements upon a partial observation of the other elements, since the principal needs to fully *commit* to a mechanism and announce the mechanism description in advance.

However, for delegated choice, one can impose more complicated *combinatorial structure* on the mechanism's decisions. For instance, one can impose a bundle-wise threshold such that the mechanism will accept a set of elements only if its total utility is at least the value of such threshold. Note, however, that this is not possible for prophet inequality as an algorithm needs to *irrevocably* decide whether or not to accept each arriving element without knowing the future element's values. On the other hand, for example, a prophet inequality algorithm can adjust its decision on the fifth element after accepting the

[2]When we say an algorithm, it refers to a prophet inequality algorithm, and a mechanism refers to the mechanism for delegated choice.

first two elements while rejecting the following two elements, whereas a delegated choice mechanism cannot adjust its description on its fifth element in such an adaptive manner. We finally remark that our mechanism for downward-closed constraints largely exploits the fact that the mechanism can impose a combinatorial structure on its decision, which indeed cannot be implemented as an algorithm for prophet inequality as mentioned above.

As conceptual contributions, our work first reveals the *intrinsic distinction* between delegated choice and prophet inequality, presenting an almost complete understanding of the differences and similarities between the two problems. Further, our work emphasizes the importance of several *variants of non-adaptive* algorithms for matroid prophet inequality, which have received less or no attention in the literature, to the extent of our knowledge. Lastly, our result on the negligible power of knowing the arrival order distribution for non-adaptive algorithms in prophet inequality complements the recent line of work on revealing the gains from having (partial) information on the arrival order [48, 49].

1.2 Our results and techniques

Before getting into the results, we briefly introduce some notation. We mainly consider two types of mechanisms: (i) agent-oblivious mechanisms in which the principal does not have any prior information regarding the distribution of the agent's utility on each solution, and (ii) agent-aware mechanisms where the principal knows the agent's utility distribution in advance and can observe the agent's utility of the proposed elements. Also, we consider several types of adversaries for the prophet inequality. An order-oblivious algorithm does not have any information regarding the arrival order, and an order-aware algorithm knows the arrival order. For both types of algorithms, the adversary picks the worst-case arrival order. An order-hinted algorithm knows the distribution from which the

arrival orders are drawn, and the adversary chooses the worst-case distribution of the arrival orders. In particular, throughout the paper, a non-adaptive algorithm for the prophet inequality refers to the algorithm that imposes a non-adaptive threshold for each element.

In what follows, we first introduce our result for the downward-closed constraints, and then we present the deep connection between the matroid prophet inequality and the delegated choice problem with matroid constraints.[3] Afterwards, we present our results on agent-aware mechanisms. Several technical lemmas and further discussion are deferred to the appendix.

1.2.1 Agent-oblivious mechanisms for downward-closed constraints

Towards designing efficient mechanisms, the most direct approach would be to observe whether the connection by [3] carries over, since it might allow us to implement delegation mechanisms from prophet inequalities. To this end, we first briefly introduce the intuition behind the results by [3]. In the single-choice setting, they provide an exact correspondence between an oblivious stopping rule in a *continuous* version of the prophet inequality and the single-proposal mechanism in the single-choice delegation problem. Their intuition behind such correspondence is as follows. Let us plot the principal's and the agent's utility (X_i, Y_i) of each solution i in the 2-dimensional Euclidean space, and consider an instance of prophet inequality such that the random variables $W_i = X_i$ arrive in an order of decreasing Y_i. The oblivious stopping rule, or in particular non-adaptive algorithm, accepts the first element that belongs to an eligible region, *i.e.,* the earliest arriving element satisfying $X_i \geq \tau$ for some threshold τ. With the single-proposal mechanism with eligible set $X_i \geq \tau$, the agent submits a solution such that it belongs to the eligible set while maximizing the agent's utility Y_i, *i.e.,* the element satisfying $X_i \geq \tau$, while maximizing Y_i. Due to the way we couple the arrival time and the agent's

[3] Preliminaries about matroids can be found in Section 1.4 and Appendix 1.4.1.

utility, the two problems' outcomes are exactly equivalent to each other. Therefore, one fundamental question is whether such correspondence remains true beyond the single-choice setting.

In the following example, however, we observe that the connection does not hold with arbitrary set constraints.

Example 1.2.1. Define a set system $\mathcal{M} = (U, \mathcal{I})$ with the universe $U = \{1, 2, 3, 4, 5\}$, where $\mathcal{I} = \{I \subset U : |I| \leq 3\} \setminus (\{1, 2, 3\} \cup \{1, 2, 4\} \cup \{1, 2, 5\})$, *i.e.,* it is a 3-uniform matroid where three specific bases are excluded. Let the support of all the random variables be within $[0, 1]$, and consider any non-adaptive algorithm ALG for the prophet inequality. Consider a realization of random variables $(X_i, Y_i)_{i \in U}$ such that $Y_i = 1/(i + 1)$ and principal's utility satisfies that $X_1 = X_2 = 1$, and $X_i = 0.1$ for $i = 3, 4, 5$. Consider a coupling $W_i = X_i$ for $i \in U$ in the corresponding prophet inequality problem where the elements arrive in decreasing order of Y_i, and consider ALG. Since the support is limited to lie in $[0, 1]$, obviously ALG's threshold for each random variable $\tau_i \in [0, 1]$. Hence, ALG selects the first two elements W_1, W_2, and skips the remaining elements since all of the elements $3, 4$ and 5 are forbidden after selecting elements 1 and 2. On the other hand, the agent's best response given the realization is to propose $\{1, 4, 5\}$. Indeed, agent's best response yields the principal utility of 1.2, but ALG obtains at least 2. Thus, not only the agent's best response and the outcome of ALG differ, but such different outcome yields strictly lower utility for the principal as well, compared to what the algorithm has obtained in the prophet inequality.

This implies that designing delegation mechanisms from prophet inequality is hopeless for downward-closed constraints, and we need a unique device tailored to the delegated choice problem to tackle it.[4] Our intuition for downward-closed mechanisms starts from the fact that in the delegated

[4]In fact, there are provably no constant competitive algorithms for the prophet inequality with general downward-closed set systems as per [8]. Our intention is to show that even though there exist some algorithms for the prophet inequality, the correspondence no longer holds due to the inconsistency between the algorithm's choice and the agent's best response.

choice with combinatorial constraints, we can impose more combinatorial thresholds on the acceptable sets, *e.g.,* by announcing a threshold on the entire utility obtainable from the proposed set. We describe below some further technical tools used in obtaining a constant-approximate mechanism, but let us first present the theorem statement here.

Theorem 1.2.2 (Constant approximation for downward-closed constraints; Informal)**.** *Under the downward-closed set constraint, there exists a constant approximate delegation mechanism.*

To elucidate an implication of this result, we emphasize that for the prophet inequality with general downward-closed constraints, the best possible approximation factor is known to be only $\log\log(n)/\log(n)$ by [45] and [8]. Thus, there exists a *strict separation* between the downward-closed prophet inequality and the downward-closed delegated choice problems, implying that the delegated choice can go much beyond what the prophet inequality can do. Indeed, we note that our mechanism used for Theorem 1.2.2 cannot be implemented as a prophet inequality algorithm.

The main technical challenge arises from the fact that neither element-wise thresholds-only nor bundle-wise[5] thresholds-only work here. Indeed, we first observe from Example 1.2.1 that only imposing element-wise thresholds does not work. This is because if this does work, this actually can be implemented as a prophet inequality algorithm with non-adaptive thresholds, implying a constant-factor approximation for downward-closed prophet inequality, which contradicts [8]. On the other hand, imposing only bundle-wise thresholds does not work as well. One can mimic the analysis of the standard single-choice prophet inequality by considering each bundle $I \in \mathcal{I}$ as a single element in the prophet inequality. However, the negative correlation between bundle I's expected value conditioned on the bundle J being lower than the specified threshold may lead to an arbitrarily bad performance of the mechanism.

[5]Bundle here refers to a feasible set $I \in \mathcal{I}$.

Indeed, our mechanism smartly uses both features by combining the idea of (i) the median-based algorithm for single-choice prophet inequality and (ii) the core-tail decomposition technique by [50]. We refer to Section 1.5 for more intuitions. It starts from decomposing the core and tail contribution of mechanism MEC. Precisely, the tail contribution TAIL(MEC) comes from the elements whose values are larger than or equal to some large enough threshold τ_t, and the core contribution CORE(MEC) comes from the remaining elements. Our tail threshold τ_t will be defined to be α-quantile of the random variable $\max_{i \in U} X_i$ for an appropriate choice of α.

Note that the total expected utility can be written as $\mathbb{E}[\,\text{MEC}\,] = \text{CORE}(\text{MEC}) + \text{TAIL}(\text{MEC})$. If $\text{CORE}(\text{OPT}) \geq \text{TAIL}(\text{OPT})$ where OPT denotes the optimal feasible set in hindsight, we simply lower bound the CORE(OPT) in our mechanism, and otherwise, we lower bound TAIL(OPT). Suppose that $\text{CORE}(\text{OPT}) < \text{TAIL}(\text{OPT})$. Due to our choice of τ_t, we can lower bound $\mathbb{E}[\,\text{MEC}\,]$ by an α fraction of the summation of each element's tail contribution. This is because our choice of τ_t guarantees that only a single element will have value above τ_t with probability at least α. Further, our choice of τ_t will eventually imply that conditioned on the event that this single element's value is at least τ_t, its conditional expectation will be a good approximation of the optimal benchmark. Further, $\mathbb{E}[\,\text{OPT}\,]$ can be upper bounded by $2 \cdot \text{TAIL}(\text{OPT})$, which itself can be bounded above by the summation of the tail contribution of every element; this leads to a constant approximation factor overall.

On the other hand, if $\text{CORE}(\text{OPT}) \geq \text{TAIL}(\text{OPT})$, it suffices to lower bound the mechanism's performance given that all the random variables are bounded above. If the boundary τ_t is large enough so that it is bounded below by $\beta \cdot \text{CORE}(\text{MEC})$, we can argue that simply running the same mechanism as above yields a $\beta \cdot \alpha$-approximation, since τ_t is selected in a way that at least one element is at least τ_t with constant probability α. Otherwise, *i.e.,* $\tau_t < \beta \cdot \text{CORE}(\text{MEC})$, we can apply the concentration of self-bounding function by [51] to argue that with some good constant probability, the core contribution

of the maximum weight feasible set is at least a δ of its expectation. Thus, imposing the bundle-threshold $\tau_c = \delta \cdot \textsc{Core}(\textsc{Opt})$, *i.e.,* to accept the set if the total utility from the set is at least τ_c, gives us a constant factor approximation.

1.2.2 Agent-oblivious mechanisms for matroid constraints

Example 1.2.1 explicitly shows that the correspondence does not hold for arbitrary set constraints. Therefore, the fundamental question one can ask is to identify the family of set constraints under which the correspondence remains true. To elaborate on our results, we need to introduce an (informal) definition of an implementable algorithm. Informally, we say that an algorithm for the prophet inequality is *implementable*, if it can be used to construct a delegation mechanism with the same approximation factor for the delegated choice problem.[6][7]

Our first main result can be stated as follows.

Theorem 1.2.3 (Non-adaptive algorithms; Informal)**.** *Any non-adaptive algorithm for the matroid prophet inequality is implementable. Specifically, under the k-uniform matroid constraint, there exists a matroid delegation mechanism with an approximation factor of $\left(1 - O(\sqrt{\frac{\log k}{k}})\right)$.*[8]

The main intuition of how we obtain this result is as follows. We start with the coupling technique between the prophet inequality and delegated choice problem, *i.e.,* $W_i = X_i$ and the random

[6] A formal definition of the implementability can be found in Definition 1.4.1.

[7] We mostly consider the setting in which the principal's and agent's utility X_i and Y_i are independent for $i \in [n]$, and consider prophet inequality algorithms against an offline adversary who determines the arrival order before realization of any random variable. However, we emphasize that this is for sake of simplifying our exposition and to focus on main implications and takeaways of our results. Indeed, all the implementability results to be presented here carry over to the setting in which X_i and Y_i can be correlated if the prophet inequality algorithm's performance is guaranteed against an almighty adversary who observes all the realization of the random variables in advance and determines the arrival order. Detailed discussion can be found around Corollary 1.6.2.

[8] More precise statement along with exact approximation factors can be found in Theorem A.4.2 and Corollary 1.6.3. Specifically, we use *approximation factor* for delegation mechanism, and *competitive ratio* for prophet inequality algorithms to denote the corresponding approximation ratios.

variables arrive in a decreasing order of Y_i, under the matroid structure $\mathcal{M}$. Given a non-adaptive threshold τ_i for each i and given the realization of the random variables, we can identify a set of random variables that are acceptable with respect to the thresholds, *i.e.,* $X_i \geq \tau_i$. Denoting such a set by $V = \{i \in U : X_i \geq \tau_i\}$, one can observe that $\mathcal{M}|V$, *i.e.,* the set system restricted to V, is still a matroid. Then, we consider a *greedy algorithm* in the restricted matroid $\mathcal{M}|V$ to find the maximum weight basis, where the weights are given by Y_i. Since the greedy algorithm indeed finds the maximum weight basis, this implies that the agent's best response to maximize his own utility is to greedily select the elements in decreasing order of Y_i, under the restricted matroid $\mathcal{M}|V$. Therefore, this establishes the correspondence between the delegated choice and the non-adaptive algorithm in the matroid prophet inequality, since the algorithm also chooses the feasible elements, *i.e.,* those in V, with respect to the arrival order, which is the decreasing order of Y_i.

Necessity of matroid structure for implementation. As acute readers may have noticed, the main intuition behind Example 1.2.1 and Theorem 1.2.3 gives some evidence that the matroid structure might be very essential to make an algorithm implementable.

The following theorem formalizes and confirms this intuition.

Theorem 1.2.4 (Matroid necessity; Informal)**.** *Suppose that $\mathcal{M} = (U, \mathcal{I})$ is a set system with downward-closed constraint $\mathcal{I}$. Consider any non-adaptive algorithm* ALG *whose thresholds are non-negative finite values for every problem instance. If* ALG *under $\mathcal{M}$ is implementable to its induced mechanism* MEC*, then $\mathcal{M}$ is a matroid.*

The formal definition of an induced mechanism can be found in Section 1.7. En route to this result, we need a slight variant of the well-known result which states the necessity of being matroid when the greedy algorithm is optimal for any non-negative integer-valued weight functions, formally

stated in Lemma A.4.3. More formally, equipped with this lemma, the proof of Theorem 1.2.4 proceeds by considering an exponentially decaying discrete distribution such that its probability density function is defined as $f(i) = 2^{-i-1}$. Then, we essentially consider a problem instance of the delegated choice in which all the random variables follow this distribution. Our construction of the problem instance guarantees that every non-negative integer vector can be realized as the agent's utility with strictly positive probability, and we apply the aforementioned lemma to obtain the final result.

Implementation of broader class of non-adaptive algorithms. With purely non-adaptive algorithms for matroid prophet inequality, however, it is hopeless to obtain constant factor approximations, as [52] provided the upper bound of $\Omega(\log\log n/\log n)$ on the approximation ratio of any non-adaptive algorithm. However, we overcome this impossibility result, by observing that a broader class of algorithms, defined as *matroid-constrained* non-adaptive algorithms,[9] can be implemented as well, using a simple variant of Theorem 1.2.3.

Proposition 1.2.5 (Matroid-constrained non-adaptive algorithms; Informal)**.** *Any matroid-constrained non-adaptive algorithm for the matroid prophet inequality is implementable. Specifically, there exists a* 3*-approximate agent-oblivious mechanism for the matroid delegated choice on graphic matroids, and a* 4*-approximate mechanism for this problem on general matroids.*

It is worth mentioning, that this result highlights the importance of designing matroid-constrained non-adaptive algorithms for matroid prophet inequality, which previously had no further implications other than its usefulness for the prophet inequality itself.

Note further that the super-constant upper bound of $\Omega(\log\log n/\log n)$ by [52] applies to purely non-

[9] A constrained non-adaptive algorithm imposes a sub-constraint $\mathcal{J} \subset \mathcal{I}$ upon the thresholds and accepts the element with respect to $\mathcal{J}$ given the threshold condition. In particular, we focus on the matroid-constrained non-adaptive algorithm whose sub-constraint $\mathcal{J}$ itself constitutes a matroid over the restricted set of elements. Formal definitions can be found in Section 1.4.2.

adaptive algorithms in which their decision at the boundary, *i.e.,* when $X_i = \tau_i$, is always deterministic. In fact, we observe that the reduction in Theorem 1.2.3 can be combined with perturbation techniques to construct non-adaptive algorithms whose decisions at the boundary are randomized, which we call the *boundary-randomized* non-adaptive algorithms.

Our implementability result can be stated informally as follows.

Theorem 1.2.6 (Boundary-randomized non-adaptive algorithms; Informal)**.** *For any $\varepsilon > 0$, any boundary-randomized non-adaptive α-competitive algorithm can be implemented to an $(\alpha-\varepsilon)$-approximate mechanism for the matroid delegated choice.*[10]

In particular, implementing a boundary-randomized non-adaptive algorithm is not straightforward since it requires the selfish agent to break the tie in a precisely desired manner. This is straightforward in the prophet inequality, since the algorithm can break the tie however it likes given the online arrival of the random variables. In the delegated choice problem, however, the random variables in fact do not arrive in an online manner, and we cannot directly enforce the agent to best respond with respect to our desired tie breaking rule. Thus we require a reduction that allows the principal to affect how the agent breaks ties exactly in the way she desires.

We achieve this by designing a mechanism that slightly perturbs the utility distribution of the given problem instance. Then, we also apply the same perturbation to the utility of random variables that the agent has proposed in examining the proposed set. Combined with several technical lemmas stating that the outcome of the perturbed problem instance can be made sufficiently close to the original problem instance, we can now apply Theorem 1.2.3 to conclude the proof.

[10]We note that a slight variant of the proof of Theorem 1.2.4 ensures that both the matroid-constrained non-adaptive algorithm and boundary-randomized non-adaptive algorithm are subject to the matroid necessity result in Theorem 1.2.4, but we omit the proof as it is a cumbersome extension.

1.2.3 Agent-aware mechanisms

Finally, we investigate the power of agent-aware mechanisms over agent-oblivious mechanisms. If the principal knows the agent's utility distribution, she can exploit this distribution to run a smarter mechanism. Indeed, by slightly modifying the proof of Theorem 1.2.3 to exploit the knowledge of the agent's distribution, we find that a stronger class of algorithms can be implemented as agent-aware mechanisms.

Proposition 1.2.7 (Implementing order-hinted algorithms; Informal)**.** *Any α-competitive order-hinted non-adaptive algorithm for matroid prophet inequality implements an α-approximate agent-aware mechanism for the matroid delegated choice.*

As a corollary of independent interest, we find that a non-adaptive algorithm for *matroid prophet secretary*[11] can be implemented to an agent-aware mechanism under an ex-ante indifferent setting, thereby suggesting the importance of designing non-adaptive matroid prophet secretary algorithms, which has not been studied before. We refer to Corollary A.2.3 for more details.

This gives some evidence that the agent-aware mechanisms might be strictly more efficient than the agent-oblivious mechanisms. However, the following theorem proves that order-hinted algorithms are *not strictly better* than order-oblivious algorithms, thereby suggesting that the agent-aware mechanism induced from an order-hinted algorithm is *not strictly better* than the agent-oblivious mechanism induced from an order-oblivious mechanism.

Theorem 1.2.8 (Power of knowing the arrival distribution; Informal)**.** *Let* ALG *be any α-competitive order-hinted non-adaptive algorithm for prophet inequality with set system $\mathcal{M}$. Then, for any $\varepsilon > 0$,*

[11] In the prophet secretary problem, the arrival order is given uniformly at random.

there exists an $(\alpha - \varepsilon)$-competitive order-oblivious non-adaptive algorithm for prophet inequality with set system $\mathcal{M}$.

To the best of our knowledge, there has been no such result regarding the power of order-hinted algorithm for the prophet inequality. Our result highlights that one cannot hope for better competitive ratio with the information on the arrival distribution, with respect to the worst-case distribution of the arrival orders.

Intuitively, to prove this result, we consider a game between an algorithm designer and an adversary who chooses the arrival order. The action space of the algorithm designer is a set of all possible deterministic non-adaptive algorithms, while the action space of the adversary is the set of permutations over the universe U. The algorithm designer tries to maximize the expected output of the algorithm where the expectation is taken over the randomness in the weights of the elements, while the adversary chooses the arrival order to minimize it. A mixed strategy of the algorithm designer is a randomized algorithm which randomly outputs the non-adaptive algorithms. Moreover, a mixed strategy of the adversary is the distribution over the arrival orders.

Now, if the algorithm designer moves first, then this corresponds to the setting in which the algorithm designer designs an order-oblivious non-adaptive algorithm, and then the adversary picks the worst-case arrival order. On the other hand, if the adversary moves first, this is the case that the arrival order is known, and thus the algorithm designer chooses an order-hinted algorithm to maximizes its utility. Thus, if one can essentially apply the minimax theorem, formally presented in Lemma A.3.2, it implies that there exists no benefit of order-hinted algorithms. The subtlety, however, lies in the fact that the collection of all possible non-adaptive algorithms is not finite, thus we cannot apply the minimax theorem directly. Instead, we construct a finite approximate class of algorithms to bound the

loss from the entire space of non-adaptive algorithms within a small factor, and apply the minimax theorem over this restricted class of algorithms.

1.3 Model

In the *delegated choice with combinatorial constraints*, a set of elements $U = \{1, 2, \ldots, n\}$ are given. A principal and an agent have utility X_i and Y_i, respectively, for each element $i \in U$. Random variables X_i and Y_i are supported on a set of non-negative real values Φ_i and Ψ_i for $i \in U$, respectively. In the independent utility setting, all the random variables $(X_i)_{i \in U}$ and $(Y_i)_{i \in U}$ are *mutually independent*. In the correlated setting, X_i and Y_i can be correlated with each other for each $i \in U$. Unless specified explicitly, our results hold under the general correlated setting. We write X and Y to denote the random vectors $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_n)$, respectively. We also write x to denote the realization of a random vector X such that $x = (x_1, x_2, \ldots, x_n)$ and x_i denotes the realization of X_i, and similarly for y and y_i. Further, there exists a set of feasible allocations $\mathcal{I}$ defined over U, that is $\mathcal{I} \subseteq 2^U$,[12] which is publicly available to both the principal and the agent. We let $\mathcal{M} = (U, \mathcal{I})$ be a *set system* defined over U. A *problem instance* of the delegated choice problem is completely specified by the tuple $(\mathcal{M}, X, Y)$. If the underlying set system $\mathcal{M}$ is a matroid, this is called a *matroid delegated choice* problem.

The principal wants to find a maximal feasible set $I \in \mathcal{I}$ that maximizes his utility $\sum_{i \in I} X_i$. However, there exists an *information asymmetry* such that the principal cannot observe the realization of random variables, but only the agent can do. As a result, the principal *delegates* the role of selecting the feasible set to the agent. Technically, the principal only knows the distribution of (X, Y), whereas the agent observes the *realization* of (X, Y). The agent is rational *expected utility maximizer*, and thus

[12] 2^U denotes the power set over U.

proposes a feasible set in order to maximize his own expected payoff given the mechanism. If the mechanism's final allocation is given by $I \in \mathcal{I}$, the eventual utility of the principal is $\sum_{i \in I} X_i$, and that of the agent is $\sum_{i \in I} Y_i$, *i.e.,* utility functions are additive. We often write X_I to denote $\sum_{i \in I} X_i$ and similarly for Y_I, unless confusion arises.

If there exists no condition on an acceptable set of elements, the agent will simply solve the optimization problem $\max_{I \in \mathcal{I}} \sum_{i \in I} Y_i$ after observing the realization, and propose the solution of it. The principal's utility is then $\sum_{i \in I} X_i$, which can be arbitrarily bad if X_i is largely misaligned with Y_i. Thus, the principal faces a *moral hazard* of the misaligned agent. To optimize the expected utility, the principal commits to an *eligible collection* as in [3].[13] Formally, an eligible collection is a function $\mathcal{E} : \mathcal{I} \times \left(\bigtimes_{i \in U} (\Phi_i, \Psi_i)\right) \mapsto \{0, 1\}$ that maps the agent's proposed set $I \in \mathcal{I}$ and utility vector $\big((X_i, Y_i)\big)_{i \in U}$ to the binary signal of whether the principal accepts the proposed set or not. For example, if the agent proposes $\{2, 3\}$ among $\{1, 2, 3\}$, then $\mathcal{E}$ decides whether or not to accept $I = \{2, 3\}$ after observing the realization of $\big((0, 0), (X_2, Y_2), (X_3, Y_3)\big)$. Note that since element 1 is not contained in the proposed set, the mechanism cannot observe its quantity and thus is simply masked by pair of zeros. This is called a *single-proposal mechanism*, and we denote by $\textsc{Mec}_{\mathcal{E}}$ the mechanism that posts an eligible collection $\mathcal{E}$.

Given a problem instance $(\mathcal{M}, X, Y)$, if the principal knows the realization of the random variables and solves the task by herself, then the expected optimal value in hindsight will be the following:

$$\mathbb{E}\left[\,\textsc{Opt}(\mathcal{M}, X, Y)\,\right] = \mathbb{E}\left[\max_{I \in \mathcal{I}} \sum_{i \in I} X_i\right]. \tag{1.1}$$

[13] This is defined as a *single-proposal mechanism* in [3]. In fact, one can consider a more generalized class of mechanisms, but due to a revelation principle style of argument, it is without loss of generality to only consider single-proposal mechanism in our setting. Details can be found in Proposition 1.4.2.

However, since it is the agent who proposes the set, the principal's utility depends on the agent's strategy. Formally, the agent's strategy is a mapping $\sigma : \left(\bigtimes_{i\in U}(\Phi_i, \Psi_i)\right) \mapsto \mathcal{I}$, which maps the observed utility $(x_i, y_i)_{i\in U} \in \bigtimes_{i\in U}(\Phi_i, \Psi_i)$ to the proposed set $I \in \mathcal{I}$. Given that the principal commits to an eligible collection $\mathcal{E}$, the agent plays a *best-response* strategy that maximizes his own utility. Given a problem instance $(\mathcal{M}, X, Y)$, we denote the agent's best-response given a mechanism $\textsc{Mec}_{\mathcal{E}}$ by $\mathcal{B}(\textsc{Mec}_{\mathcal{E}}; \mathcal{M}, X, Y)$. Since it maximizes the agent's utility, we also have

$$\mathcal{B}(\textsc{Mec}_{\mathcal{E}}; \mathcal{M}, X, Y) = \operatorname*{argmax}_{I\in\mathcal{I}} \left\{ \sum_{i\in I} Y_i \quad \text{subject to} \quad \mathcal{E}\left(I, (X_i, Y_i)_{i\in I}\right) = 1 \right\}. \tag{1.2}$$

We often write $\mathcal{B}(\textsc{Mec}_{\mathcal{E}}) = \mathcal{B}(\textsc{Mec}_{\mathcal{E}}; \mathcal{M}, X, Y)$ if the problem instance is clear from the context. A mechanism $\textsc{Mec}$ is α-approximate if the principal's expected utility under $\textsc{Mec}$ is at least $\alpha \cdot \mathbb{E}\left[\,\textsc{Opt}(\mathcal{M}, X, Y)\,\right]$, for some $\alpha \in [0, 1]$, for every problem instance $(\mathcal{M}, X, Y)$. We write $\textsc{Mec}(\mathcal{M}, X, Y)$ to denote the expected utility from running $\textsc{Mec}$ given the problem instance $(\mathcal{M}, X, Y)$. More formally,

$$\mathbb{E}\left[\,\textsc{Mec}_{\mathcal{E}}(\mathcal{M}, X, Y))\,\right] = \mathbb{E}\left[\sum_{i\in\mathcal{B}(\textsc{Mec}_{\mathcal{E}})} X_i\right]. \tag{1.3}$$

We simply write $\mathbb{E}\left[\,\textsc{Opt}\,\right]$ and $\mathbb{E}\left[\,\textsc{Mec}_{\mathcal{E}}\,\right]$ to denote (1.1) and (1.3), if the problem instance is clear from the context. If there are multiple feasible sets that optimize (1.2), we assume that the agent breaks the tie arbitrarily.

We consider a *Bayesian mechanism* in which the principal knows the distribution of X, and tries to compete with the optimal value in hindsight, defined as (1.1). We say that the principal is *agent-aware* if she knows the distribution of the agent's utility and can observe the agent's realized

utility for the proposed set, and *agent-oblivious* if it has no information regarding the agent's utility, nor cannot observe the agent's utility of the proposed set. For both cases, the approximation factors are measured based on the worst case problem instance $(\mathcal{M}, X, Y)$, *i.e.,* the underlying set constraint and the distributions of the random variables.

1.4 Preliminaries

1.4.1 Matroids

A set system is a pair $\mathcal{M} = (U, \mathcal{I})$ such that U is a finite set of *elements*, called a *ground set*, and $\mathcal{I} \subseteq 2^U$ is a family of *feasible* sets. A *downward-closed* set system satisfies that $\emptyset \in \mathcal{I}$ and if $T \in \mathcal{I}$ then $S \in \mathcal{I}$ for every $S \subseteq T$. A *matroid* $\mathcal{M}$ is a downward-closed set system such that $\mathcal{I}$ satisfies the following *augmentation* property: if $A, B \in \mathcal{I}$ and $|A| > |B|$, then there exists $x \in A \setminus B$ such that $B \cup \{x\} \in \mathcal{I}$. If $\mathcal{M}$ is a matroid, any $I \in \mathcal{I}$ is called an independent set, and a basis is a maximal independent set. Denote the set of all bases by $\mathcal{B}$. We further define several standard notions. Given a $V \subseteq U$, the *restriction* of $\mathcal{M}$ to V is $\mathcal{M} \mid V = (V, \mathcal{I}|V)$ where $\mathcal{I}|V = \{I \cap V : I \in \mathcal{I}\}$. Given a $V \subseteq U$, the *contraction* of $\mathcal{M}$ by V, written $\mathcal{M}/V$, is the set system over $U \setminus V$ whose feasible sets are the sets whose union with V is independent in $\mathcal{M}$. A *matroid minor* of the matroid $\mathcal{M}$ is another matroid obtained from $\mathcal{M}$ by a sequence of restrictions and contractions. For a subcollection $\mathcal{J} \subset \mathcal{I}$, we define *induced set system* be a $\mathcal{M}|\mathcal{J} = (U|\mathcal{J}, \mathcal{J})$ where $U|\mathcal{J} = \{i \in U : \exists J \in \mathcal{J}, i \in J\}$. If $\mathcal{M}|\mathcal{J}$ constitutes a matroid, we call it an *induced matroid*. We write $\mathcal{I}(\mathcal{M})$ to denote the collection of independent sets equipped to $\mathcal{M}$. A *rank* of a set $V \subseteq U$ for matroid, denoted by $\text{rank}(V)$, is the size of a largest independent set contained in V. A *matroid polytope* is defined as $\mathcal{P}_{\mathcal{M}} = \{p \in \mathbb{R}_{\geq 0}^{|U|} | \forall S \in 2^U, \sum_{i \in S} p_i \leq \text{rank}(S)\}$, *i.e.,* the convex hull of the characteristic vectors of the bases of $\mathcal{M}$. One can

define a *set polytope* for arbitrary set system $\mathcal{M}$ as a convex hull of all characteristic vectors of feasible sets, *i.e.,* $P_{\mathcal{M}} = \text{conv}(\{\mathbf{1}_I | I \in \mathcal{I}\})$, too. Unless stated explicitly, the universe is simply $U = [n]$.

1.4.2 Prophet inequality

The *prophet inequality* problem is defined over a set system $\mathcal{M} = (U, \mathcal{I})$ with $U = [n]$. Each $i \in U$ has a random variable W_i with distribution F_i supported on a set of non-negative real numbers, which denotes element i's *random weight*. Let $\mathcal{F}_i$ be the collection of all possible distributions F_i on $\mathbb{R}_{\geq 0}$, and $\mathcal{F} = \bigtimes_{i \in U} \mathcal{F}_i$. Each element i is further equipped with a random *arrival time* $T_i \in \mathbb{R}_{\geq 0}$, and the elements arrive in increasing order of the arrival time.[14] Each T_i is equipped with distribution G_i supported on non-negative real numbers. Let $\mathcal{G}_i$ be the collection of all possible distributions G_i on $\mathbb{R}_{\geq 0}$, and $\mathcal{G} = \bigtimes_{i \in U} \mathcal{G}_i$. Further, we use Γ to denote the collection of all possible deterministic distributions on $\mathbb{R}^n_{\geq 0}$, *i.e.,* $\gamma \in \Gamma$ is completely specified by a n-dimensional non-negative real vector.[15] We write W to denote the random vector $(W_1, \ldots, W_n)$, and similarly for T. Overall, a *problem instance* of prophet inequality is given by $(\mathcal{M}, W, T)$.

An online algorithm for prophet inequality is an algorithm such that at every round, if element i arrives, it irrevocably decides whether or not to accept i, upon observing the realization of W_i, subject to the feasibility constraint $\mathcal{I}$, *i.e.,* the selected set I must be feasible $I \in \mathcal{I}$. We write W_I to denote $\sum_{i \in I} W_i$. Given a problem instance $(\mathcal{M}, W, T)$, we denote the resulting allocation set given an algorithm ALG by $\mathcal{A}(\text{ALG}; \mathcal{M}, W, T)$. We often write $\mathcal{A}(\text{ALG})$ if the problem instance is clear from the context. The algorithm competes with the maximal feasible set with maximal weight in hindsight

[14] In fact, only the ordering of the arrival times matters, and their magnitudes are irrelevant.
[15] This will be used to define the competitive ratio of an order-oblivious algorithm shortly.

defined as follows:

$$\mathbb{E}\left[\,\textsc{Opt}(\mathcal{M}, W, T)\,\right] = \mathbb{E}\left[\max_{I\in\mathcal{I}} \sum_{i\in I} W_i\right]. \tag{1.4}$$

Similarly, an algorithm ALG's performance given the problem instance is defined as follows:

$$\mathbb{E}\left[\,\textsc{Alg}(\mathcal{M}, W, T)\,\right] = \mathbb{E}_{W\sim F, T\sim G}\left[\sum_{i\in\mathcal{A}(\textsc{Alg}, W, T)} W_i\right]. \tag{1.5}$$

We often write $\mathbb{E}\left[\,\textsc{Alg}\,\right]$ and $\mathbb{E}\left[\,\textsc{Opt}\,\right]$ for (1.4) and (1.5) if the problem instance is clear from the context. Throughout the paper, we write ALG to denote an algorithm for the prophet inequality problem, and MEC to denote a mechanism for the delegated choice problem.

Arrival order and competitive ratio. We consider several types of algorithms based on the algorithm's prior knowledge on the arrival order T. An *order-aware* algorithm is an algorithm that knows the realization $t = (t_1, \ldots, t_n)$ of the arrival orders T exactly. An *order-oblivious* algorithm does not have any information regarding T, nor its distribution. For both types of algorithms, a *competitive ratio* of an algorithm ALG is defined as follows:

$$\textsc{Comp}(\textsc{Alg}) = \min_{F\in\mathcal{F}, G\in\Gamma} \frac{\mathbb{E}_{W\sim F, T\sim G}\left[\textsc{Alg}(\mathcal{M}, W, T)\right]}{\mathbb{E}_{W\sim F, T\sim G}\left[\textsc{Opt}(\mathcal{M}, W, T)\right]},$$

where the expectation is taken over the randomness of W as well as the random bits of the algorithm. Note further that T does not have any randomness in this case.

Further, an *order-hinted* algorithm is the one that only knows the distribution $G \sim \mathcal{G}$ from which the arrival orders are sampled. In other words, it knows that the arrival order is sampled from a

distribution $G \sim \mathcal{G}$, but does not exactly know their realization. In this case, ALG's competitive ratio is defined over the worst-case possible distribution and distribution of the random weights:

$$\textsc{Comp}(\textsc{Alg}) = \min_{F \in \mathcal{F}, G \in \mathcal{G}} \frac{\mathbb{E}_{W \sim F, T \sim G}\left[\textsc{Alg}(\mathcal{M}, W, T)\right]}{\mathbb{E}_{W \sim F, T \sim G}\left[\textsc{Opt}(\mathcal{M}, W, T)\right]},$$

where the expectation is taken over randomness of W, T, and the algorithm's random bits.

We say that an algorithm is α-*competitive* if it's competitive ratio is at least α.

Finally, in the matroid prophet secretary problem, the arrival order is selected uniformly at random, and the algorithm knows this information. Let UNIFORM be the uniform distribution over $[|U|]$. Then, the competitive ratio is defined as

$$\textsc{Comp}(\textsc{Alg}) = \min_{F \in \mathcal{F}} \frac{\mathbb{E}_{W \sim F, T \sim \textsc{Uniform}}\left[\textsc{Alg}(\mathcal{M}, W, T)\right]}{\mathbb{E}_{W \sim F, T \sim \textsc{Uniform}}\left[\textsc{Opt}(\mathcal{M}, W, T)\right]},$$

where the expectation is taken over the randomness of W, T, and random bits of ALG.

We finally note that in all cases, the algorithm's competitive ratio is defined with respect to the *offline adversary*, which determines the problem instance before the algorithm starts. However, our results easily extend to the *almighty adversary* who can determine the arrival order after observing the realization of the random variables.

Algorithm classes. We introduce here several classes of algorithms. First, throughout the paper, we focus on a class of *threshold-based*[16] algorithms, such that if element i arrives, upon having selected a set A, the algorithm accepts the current element i if and only if $W_i \geq \tau_i$ for some nonnegative real number τ_i and $A \cup i \in \mathcal{I}$.[17] An *adaptive* algorithm adaptively changes τ_i based on the history and the

[16] In fact every reasonable class of algorithms, so-called *monotone* algorithms, can be completely defined as threshold-based algorithm. We refer to [44] for more detailed discussion.

[17] We write $A \cup i$ to denote $A \cup \{i\}$ for element i and set A.

currently selected set, while a non-adaptive algorithm determines all the thresholds $\tau = (\tau_i)_{i\in[n]}$ before any element arrives. A *constrained* non-adaptive algorithm is a non-adaptive algorithm such that upon the thresholds τ, it specifies a sub-constraint on a feasible set $\mathcal{J} \subseteq \mathcal{I}$ so that the algorithm only accepts elements that make the selected set feasible with respect to $\mathcal{J}$. A *matroid-constrained* non-adaptive algorithm is a constrained non-adaptive algorithm where the induced set system $\mathcal{M}|\mathcal{J}$ is a matroid. A *randomized* non-adaptive algorithm is a non-adaptive algorithm which randomly selects the thresholds. A *boundary-randomized* non-adaptive algorithm is a non-adaptive algorithm with thresholds τ and probability vector $p = (p_1, \dots, p_n) \in [0,1]^n$ such that if the tie happens, *i.e.,* $X_i = \tau_i$, it accepts element i with probability p_i if its addition gives a feasible soliution subject to set constraint. A *static* algorithm is a non-adaptive algorithm with a single threshold $\tau \in \mathbb{R}_{\geq 0}$ such that it accepts element i whenever $W_i \geq \tau$ and its addition preserves feasibility with respect to $\mathcal{I}$.

1.4.3 Implementation

Many parts of our main results deal with converting non-adaptive algorithms in prophet inequality to delegation mechanisms. We define here a precise notion of the implementable algorithm. We write $X \sim Y$ for random variables X and Y if their realization is coupled to be the same.

Definition 1.4.1 (Implementation)**.** Given a set system $\mathcal{M} = (U, \mathcal{I})$, consider a problem instance $(\mathcal{M}, X, Y)$ of delegated choice with combinatorial constraints. Consider a problem instance $(\mathcal{M}, W, T)$ of prophet inequality such that $W \sim X$ and $T \sim e^{-Y}$. That is, the random variables are coupled so that the realization of W_i coincides that of X_i and T_i coincides that of e^{-Y_i} for every $i \in U$. For some

mechanism MEC, we say that ALG *implements* MEC if the following holds everywhere

$$\mathcal{A}(\text{ALG}; \mathcal{M}, W, T) = \mathcal{B}(\text{MEC}; \mathcal{M}, X, Y).$$

That is, their outcomes are the same in a point-wise manner for every coupled realization of W, T, X, Y. We often say that ALG is implementable to MEC or MEC is implementable by ALG. We say that the coupled problem instances $(\mathcal{M}, X, Y)$ and $(\mathcal{M}, W, T)$ are associated with each other.

It's worth noting that in the 1-uniform matroid setting, [3] show that every oblivious stopping rule is implementable to some mechanisms, thereby constructing several constant approximate mechanisms. To see more intuition behind our definition of the implementability, we kindly ask the readers to recap the correspondence explained in Section 1.2.2.

Recall that our definition of eligible set in fact does not subsume every reasonable class of mechanisms. For example, one may consider a mechanism which requires the agent to propose at least two eligible sets, and picks one of them based on a pre-committed rule. More generally, one may consider a signaling mechanism in which the agent sends a generic signal about the realization of the random variables, and the principal selects a feasible set based on the information acquired from the signal. We provide a revelation style of argument suggesting that it is without loss of generality to focus on the class of single-proposal mechanisms. To this end, we first define a notion of interim outcome. Given a realization (x, y) of random vectors (X, Y), an interim outcome of MEC is the outcome of the mechanism given the realization, denoted by $g(\text{MEC}; \mathcal{M}, x, y) \in \mathcal{I} \times \left(\bigtimes_{i \in U} (\Phi_i, \Psi_i)\right)$.

Proposition 1.4.2 (Revelation principle)**.** *Given any instance $(\mathcal{M}, X, Y)$ of delegated choice with combinatorial constraints, for any mechanism* MEC*, there exists a single-proposal mechanism* MEC$'$ *such that their interim outcomes are equivalent,* i.e., *given arbitrary realization (x, y) of (X, Y), we*

have that $g(\textsc{Mec}; \mathcal{M}, x, y) = g(\textsc{Mec}'; \mathcal{M}, x, y)$.

1.5 Downward-closed constraint

In this section, we prove Theorem 1.2.2. Mainly, our intuition starts from a simple observation such that in the delegated choice, we can impose more combinatorial constraints on the acceptable sets, *e.g.,* announcing a threshold on the entire utility of the proposed set.

Why straightforward extension fails. Before presenting the main proof, let us first try whether the standard analysis of prophet inequality would somehow carry over. We shall briefly recap the analysis of the single-choice prophet inequality by [6]. Intuitively, we want a constant probability on the event that once the element i arrives, the element i can be selected since no other elements are selected before. To this end, let us select the threshold τ so that

$$\Pr\left[\,\text{no elements are selected}\,\right] = \Pr\left[\,\text{all the elements' values are less than } \tau\,\right] = \alpha, \tag{1.6}$$

for some constant α which will be defined shortly. Then, we have $\Pr\left[\,\text{element } i \text{ can be selected}\,\right] \geq \alpha$ since this probability is at least LHS of (1.6). For simplicity, assume $U = [n]$ and the elements arrive in a sequential order of $1, 2, \ldots, n$. Let E be the event that $X_i < \tau$ for every $i \in U$. Then, we can decompose what the algorithm obtains as follows, $\mathbb{E}\left[\,\textsc{Alg}\,\right]$ can be written as

$$\begin{aligned}
\mathbb{E}\left[\,\textsc{Alg}\,\right] &= \sum_{i \in U} \Pr\left[\,X_j < \tau, \forall j < i\,\right] \Pr\left[\,X_i > \tau \mid X_j < \tau, \forall j < i\,\right] \cdot \mathbb{E}\left[\,X_i \mid X_i > \tau, X_j < \tau, \forall j < i\,\right] \\
&\geq \sum_{i \in U} \Pr\left[\,X_j < \tau, \forall j < i\,\right] \Pr\left[\,X_i > \tau\,\right] \cdot \mathbb{E}\left[\,\tau + (X_i - \tau)^+ \mid X_i > \tau, X_j < \tau, \forall j < i\,\right].
\end{aligned}$$

Note, however, that $\sum_{i \in U} \Pr\left[\,X_j < \tau, j < i\,\right] \Pr\left[\,X_i > \tau\,\right] = \Pr\left[\,E^c\,\right]$. Thus, we can decompose the

above as follows

$$\begin{aligned}\mathbb{E}\left[\,\textsc{Alg}\,\right] &\geq \sum_{i\in U} \mathbb{E}\left[\,(X_i-\tau)^+ \mid X_j<\tau, \forall j<i\,\right]\cdot \Pr\left[\,X_j<\tau,\forall j<i\,\right] + \tau\cdot\Pr\left[\,E^c\,\right]\\ &= \sum_{i\in U} \mathbb{E}\left[\,(X_i-\tau)^+\,\right]\cdot \Pr\left[\,X_j<\tau,\forall j<i\,\right] + \tau\cdot\Pr\left[\,E^c\,\right]\\ &\geq \sum_{i\in U} \mathbb{E}\left[\,(X_i-\tau)^+\,\right]\Pr\left[\,E\,\right] + \tau\cdot\Pr\left[\,E^c\,\right].\end{aligned}$$

On the other hand, we have

$$\mathbb{E}\left[\,\textsc{Opt}\,\right] \leq \tau + \mathbb{E}\left[\,(X_{\max}-\tau)^+\,\right] \leq \tau + \sum_{i\in U}\mathbb{E}\left[\,(X_i-\tau)^+\,\right],$$

where $X_{\max} = \max_{i\in U} X_i$. If we choose $\alpha = 1/2$, then our choice of τ guarantees $\Pr\left[\,E\,\right] = \Pr\left[\,E^c\,\right]$, which implies that our algorithm is $1/2$-competitive.

Now, let us slightly tweak the analysis to obtain combinatorial version. Consider a mechanism with *bundle-threshold* τ such that it accepts the set I if $\sum_{i\in I} X_i \geq \tau$ and I is a maximal feasible set. Let $\mathcal{F}$ be the collection of maximal feasible set obtained from $\mathcal{I}$. Similar to the previous analysis, we can first extend our algorithm's gain as follows.

$$\mathbb{E}\left[\,\textsc{Alg}\,\right] \geq \sum_{I\in\mathcal{F}} \mathbb{E}\left[\,(X_I-\tau)^+ \mid X_J<\tau, \forall J\neq I, J\in\mathcal{F}\,\right]\cdot \Pr\left[\,X_J<\tau,\forall J\neq I, J\in\mathcal{F}\,\right] + \tau\cdot\Pr\left[\,E^c\,\right].$$

Then, it becomes not trivial to lower bound the conditional expectation since

$$\sum_{I\in\mathcal{F}} \mathbb{E}\left[\,(X_I-\tau)^+ \mid X_J<\tau, \forall J\neq I, J\in\mathcal{F}\,\right] \neq \sum_{I\in\mathcal{F}} \mathbb{E}\left[\,(X_I-\tau)^+\,\right], \tag{1.7}$$

due to the correlation between X_I and X_J if they share some mutual random variables $i \in I \cap J$. Indeed, the event $X_J < \tau, \forall J \neq I, J \in \mathcal{I}$ is negatively correlated with the random variable $(X_I - \tau)^+$, so the conditional expectation only decreases compared to the RHS of (1.7).

Toward efficient mechanism. Our approach builds upon the intuition above. Due to the arbitrary combinatorial structure induced by the downward-closed set constraint, it was hard to deal with the correlation in (1.7) if X_I and X_J share some mutual random variables. How can we get away from this? One obvious fact that we know about I and J is that, as long as $I \neq J$ and both of them are maximal, there exists an element $i \in I \setminus J$. Then, obviously, we can lower bound LHS of (1.7) by the conditional expectation of this single element X_i. However, the event $X_I \geq \tau$ may not give a good conditional expectation for the exact quantity $\mathbb{E}\left[\,(X_i - \tau)^+ \mid X_I \geq \tau\,\right]$, since the other elements in I may contribute more than i does.

Thus, our main idea is, instead of imposing a bundle-threshold τ which does not give a good conditional expectation for that single element i, let us go back to what we did in (1.6), and pick the same large enough threshold τ_t so that the single element's contribution gives some good constant approximation to the mechanism. This, however, does not give a good approximation if all the random variables are moderately similar. Hence we combine the idea of bundle-threshold with element-wise thresholds to obtain an efficient mechanism. More precisely, we consider two cases such that (i) given the large element-wise threshold τ_t defined as above, the contribution from *core*, *i.e.,* those from $X_i < \tau$, is dominant over the contribution from *tail*, *i.e.,* the single element i's contribution conditioned on the event $X_i \geq \tau_t$, and (ii) the other way around. In the first case, we simply lower bound the core contribution, using the fact that under the this event, we have $X_i < \tau$, *i.e.,* the random variables are essentially bounded and thus we can use concentration on it. For the latter, we follow the analysis

presented above as initial attempts.

Theorem 1.5.1 (Restatement of Theorem 1.2.2; Formal)**.** *Given downward-closed set system $\mathcal{M}$, there exists a 0.1098-approximate agent-oblivious mechanism for the delegated choice problem under $\mathcal{M}$.*

It is worth recalling that in case of arbitrary downward-closed constraint, [8] formally shows that constant approximation is not achievable for the prophet inequality. As the single-choice delegation mechanism is equivalent to the oblivious stopping rule of single-choice prophet inequality, one may expect that the multi-choice delegation entails the negative result of the multi-choice prophet inequality for general downward-closed constraint. In stark contrast, our result implies that the delegated choice with combinatorial constraints can be much better than what the non-adaptive algorithms can do in the prophet inequality with combinatorial constraints. Indeed, delegation mechanism allows more flexibility in its approach since in addition to imposing a threshold for each element, it can examine the entire proposed set as a bundle while screening out based on the entire utility from the bundle.

1.6 Matroid constraint

We present here the connection between the matroid delegated choice and the matroid prophet inequality, which yields improved approximation factor for certain classes of matroids. First, we prove that any non-adaptive algorithm for matroid prophet inequality is implementable. We then prove that boundary-randomized and matroid-constrained non-adaptive algorithms, both of which subsume the non-adaptive algorithm, are also implementable. Finally, we complement these results by showing that the matroid structure is indeed necessary for a prophet algorithm to be implementable. This shows deep connections between the prophet inequality and the delegated choice under matroid constraint.

1.6.1 Implementation from matroid prophet inequality

We start with the proof of the following theorem.

Theorem 1.6.1 (Restatement of Theorem 1.2.3; Formal)**.** *Any α-competitive non-adaptive order-oblivious algorithm against offline adversary for the matroid prophet inequality can be implemented to an α-approximate agent-oblivious mechanism for the matroid delegated choice under the independent utility setting.*

It is worth emphasizing that although this theorem guarantees the implementability under the independent utility setting, the analysis easily carries over to the *correlated utility* setting when the prophet inequality algorithm works against the *almighty adversary*. Precisely, the almighty adversary can manually change the arrival order to minimize the algorithm's performance, which means that it can arbitrarily change the agent's utility $Y_i's$ based on the realization of the principal utility X_i's, to minimize the mechanism's output. Thus, the mechanism's output is at least better than the algorithm's guarantee as the almighty adversary can also enforce the arrival order to be correlated with X_i in the same manner. This immediately implies the following corollary.

Corollary 1.6.2. *Any α-competitive non-adaptive order-oblivious algorithm against the almighty adversary for the matroid prophet inequality can be implemented to an α-approximate agent-oblivious mechanism for the matroid delegated choice under the general correlated utility setting.*

We leave the detailed proof to the readers as it is a straightforward extension of the proof of Theorem 1.6.1. We also remark that the other implementability results (Theorem 1.7.2, Proposition 1.8.1, Theorem 1.8.4) analogously extend to the prophet inequality against the almighty adversary and the mechanism under the general correlated setting. We do not include all the analogous statements to

avoid redundant arguments.[18]

To more elucidate the implications of Theorem 1.6.1, let us first consider a uniform matroid $\mathcal{M} = (U, \mathcal{I})$ with rank $k \in [|U|]$, *i.e.,* $\mathcal{I} = \{A \subset U : |A| \leq k\}$. This is often referred to as multi-unit (k-unit) setting in prophet inequality. Note that $k = 1$ corresponds to the single-item case.

The following corollary can be implied from combining the above theorem with the well-known results by [33] and [31].

Corollary 1.6.3 (Restatement of Theorem 1.2.3; Formal)**.** *Under k-uniform matroid constraint, there exists a mechanism satisfying the approximation ratio specified in Lemma A.4.2, with an asymptotic ratio of $1 - O(\sqrt{\frac{\log k}{k}})$, under the general correlated utility setting. In particular, the mechanism is completely determined by a static threshold τ such that it accepts the proposed set $I \in \mathcal{I}$ if every element $i \in I$ satisfies $X_i \geq \tau$.*

1.7 Necessity of matroid constraint for implementability

We cement the above results by showing that matroid structure is necessary to obtain the implementability, by formally proving Theorem 1.2.4. Our result for the necessity of matroid will be made clear once the induced mechanism is defined.

Definition 1.7.1 (Induced mechanism)**.** Given a set system $\mathcal{M} = (U, \mathcal{I})$, consider a problem instance $(\mathcal{M}, W, T)$ of delegated choice. Suppose non-adaptive ALG non-adaptive thresholds τ. A mechanism MEC under the set system $\mathcal{M}$ *induced by* ALG is defined as follows.

- Announces non-adaptive thresholds $\tau \in \mathbb{R}_{\geq 0}^{|U|}$.
- Accepts the proposed set I if $I \in \mathcal{I}$, and $X_e \geq \tau_e$ for every $e \in U$.

[18] It is also worth noting that our downward-closed result in Theorem 1.5.1 does not require independence as well.

Essentially, we will prove that if an algorithm can be implemented to the induced mechanism, then the underlying set is necessarily a matroid.

Theorem 1.7.2 (Restatement of Theorem 1.2.4; Formal)**.** *Suppose that $\mathcal{M} = (U, \mathcal{I})$ is a set system with downward-closed constraint $\mathcal{I}$. If an order-oblivious non-adaptive algorithm* ALG *under $\mathcal{M}$ against the offline adversary is implementable to induced mechanism $\mathcal{M}$, then $\mathcal{M}$ is a matroid.*

We note that although the theorem is proved only for the class of non-adaptive algorithms, it can easily be extended to matroid-constrained/boundary-randomized non-adaptive algorithms.

1.8 Implementation of broader class of non-adaptive algorithms

Unfortunately, however, beyond the uniform matroid, only few results are known for constant approximate non-adaptive algorithms. [53] present $1/32$-approximation non-adaptive algorithm in case of graphic matroid, motivated by multi-parameter mechanism design beyond unit-demand buyer. This was improved to $1/16$-approximation by [54] for graphic matroids on simple graphs. They provide further results on cographic, regular and laminar matroids, most of which are far much worse than $1/4$-approximation.

Indeed, [52] show that non-adaptive algorithms cannot be made constant approximation for arbitrary class of matroids, by constructing a worst-case problem with gammoid. Thus, it might be hopeless to obtain constant approximation mechanism for general matroids from non-adaptive algorithms for matroid prophet inequality.

Matroid-constrained non-adaptive algorithms. On the other hand, as one may have noticed from the proof of Theorem 1.6.1, we do not necessarily need to constrain ourselves to the class of non-adaptive algorithms. Indeed, consider a mechanism that imposes a sub-constraint on the feasible sets

upon the non-adaptive thresholds, *i.e.,* it defines a collection $\mathcal{J} \subset \mathcal{I}$, and only accepts the set that belongs to $\mathcal{J}$. If the underlying sub-constraint constitutes a matroid, a parallel of Theorem 1.6.1 shows that one can implement a mechanism from the corresponding prophet inequality algorithm, *i.e.,* matroid-constrained non-adaptive algorithm. Recall that the matroid-constrained non-adaptive algorithm is defined in Section 1.4.2.

Proposition 1.8.1 (Restatement of Proposition 1.2.5; Formal)**.** *Any α-competitive matroid-constrained order-oblivious non-adaptive algorithm for the matroid prophet inequality against the offline adversary implements an α-approximate agent-oblivious mechanism for the matroid delegated choice under the independent utility setting.*

For matroid-constrained non-adaptive algorithms, [34] first obtain a 3-competitive algorithm for the prophet inequality on graphic matroids, using non-adaptive thresholds with a partition matroid sub-constraint. Thus one can immediately obtain the following corollary.

Corollary 1.8.2 (Restatement of Proposition 1.2.5; Formal)**.** *There exists a 3-approximate agent-oblivious mechanism for graphic matroids under the correlated utility setting.*

Furthermore, we observe that the online contention resolution scheme (OCRS) by [52] also partitions the matroid into a union of matroid minors, which itself is a matroid. We observe that their method indeed constructs a matroid sub-constraint.

Corollary 1.8.3 (Restatement of Proposition 1.2.5; Formal)**.** *There exists a 4-approximate agent-oblivious mechanism for general matroids under the correlated utility setting.*

Implementing boundary-randomized non-adaptive algorithm. In fact, the worst-case problem instance by [52] assumes that the decision at the boundary case for the non-adaptive algorithm should

be deterministic, *i.e.,* if $X_i = \tau_i$ then it must accept the element i if it's feasible with respect to the set constraint. Indeed, if the algorithm's decision at the boundary can be randomized, there might be some chances that the algorithm overcomes the pessimistic upper bound. Then, one natural question is whether the boundary-randomized non-adaptive algorithm is implementable. We show that this is indeed possible.

Theorem 1.8.4 (Restatement of Theorem 1.2.6; Formal)**.** *For any $\varepsilon > 0$, any boundary-randomized α-competitive order-oblivious non-adaptive algorithm against the offline adversary can be implemented to $(\alpha - \varepsilon)$-approximate agent-oblivious mechanism for the matroid delegated choice under the independent utility setting.*

To briefly sketch the proof intuition, given a problem instance of the matroid delegated choice, we perturb the principal's utility distribution to make it continuous. Then, we run the given boundary-randomized algorithm. Thanks to Lemma A.3.4, the boundary-randomized algorithm under the perturbed instance behaves as a purely non-adaptive algorithm, so we can obtain non-adaptive thresholds. Then, we construct an induced mechanism from the obtained non-adaptive thresholds using the construction of Theorem 1.2.3. Such mechanism guarantees α-approximation for every perturbed problem instance. Then, together with Lemma A.4.4 which states that an arbitrary problem instance can be converted to a problem instance with the principal's utility distribution being continuous at the cost of a small approximation factor, the proof follows.

1.9 Agent-aware mechanisms

Thus far, we have investigated agent-oblivious mechanisms which do not assume any prior information regarding the agent's distribution. If the principal knows the agent's distribution as well, it

is fairly intuitive that the principal can do at least better than agent-oblivious mechanism by exploiting additional information. In this section, we explore the power of agent-aware mechanisms compared to the agent-oblivious mechanisms.

We start with simple observations under the matroid constraint such that an order-hinted non-adaptive algorithms can be implemented. As a special case of independent interest, this highlights the importance of non-adaptive algorithms for matroid prophet secretary, which has never been studied before. Due to the page limit, we defer relevant theorem statements to Appendix A.2.

Since the order-hinted algorithm attains much information compared to the order-oblivious algorithm for prophet inequality, one may expect that order-hinted algorithm can be strictly better than order-oblivious algorithm. Nonetheless, we obtain a pessimistic result on the power of order-hinted algorithm based on minimax-style of argument, thereby concluding that knowing the distribution does not help the algorithm for prophet inequality, in terms of the competitive ratio against worst-case distribution.

1.9.1 Power of knowing the arrival order is negligible

Meanwhile, the above results imply that one can implement stronger classes of algorithms for matroid prophet inequality if the principal knows the agent's utility distribution, which gives some evidence of the power of knowing the agent's distribution. In what follows, however, we show that knowing the arrival order does not give an advantage in terms of the worst-case competitive ratio. We remark that the competitive ratios are defined with respect to the worst-case distribution of the arrival orders (see Section 1.4.2). This implies that the agent-aware mechanism induced from an order-hinted algorithm is not strictly better than the agent-oblivious mechanism induced from an order-oblivious

mechanism.

Theorem 1.9.1 (Restatement of Theorem 1.2.8; Formal)**.** *Let* ALG *be any α-competitive order-hinted non-adaptive algorithm for the prophet inequality problem with set system $\mathcal{M}$. Then, for any $\varepsilon > 0$, there exists an $(\alpha - \varepsilon)$-competitive order-oblivious non-adaptive algorithm for the prophet inequality problem with set system $\mathcal{M}$.*

Chapter 2: Delegation with Costly Inspection

This chapter presents the main results of [9] (EC 2025).

2.1 Introduction

The fundamental assumption in the standard delegated choice problem is that the agent truthfully reports the corresponding utility of the proposed alternative. For instance, this could be an easily verifiable lie once the decision making process is over and the alternative is implemented, or it may backfire on the agent's reputation status. However, in many real-world applications, the cost of verifying the utility of the outcome might be expensive, *e.g.,* delegated decision of a governmental policy, or if the delegation happens only as a one-off interaction, so that the agent might have incentive to misreport the utility of the proposed alternative. In this context, we consider a variant of the delegated choice problem, termed *delegated choice with inspection cost* (DCIC), where the principal can further inspect the alternative at a certain cost upon the agent's signal and the agent might not truthfully report the utility of the proposed alternative.

We first observe that DCIC is a strict generalization of the recently introduced variant of the Weitzman's box problem by [10], called Pandora's box with nonobligatory inspection (PNOI) [29,55]. In the PNOI problem, there exists a set of boxes (alternatives) with a private reward for each box, the distribution of which is publicly available to the searcher. The searcher wants to commit to one of the boxes to maximize the reward. To this end, the searcher sequentially either (i) opens a box i at fixed cost c_i to observe the reward or (ii) selects a box that might be opened or closed, and finishes the searching process. Different from the standard Pandora's box problem with obligatory inspection which has a simple optimal index-based policy, PNOI problem is proven to be NP-hard [56] even if the support of each of the reward distributions is of cardinality three. Indeed, we observe that the PNOI problem reduces to an instance of DCIC problem where the cost of delegation is significantly larger than values obtainable from any alternative, which implies NP-hardness of the DCIC problem.

In this paper, we formally introduce and study the DCIC problem as extensions of both the delegated choice problem and PNOI problem, and provide the first constant approximate mechanisms for several versions. We first show that neither pure delegation policy nor pure PNOI policy can achieve a constant approximation to the optimal policy, implying that a combination of both the delegation and inspection is necessary to achieve constant approximation. Correspondingly, we consider a setting with costless delegation where the cost of the delegation is zero, and obtain a 3-approximate mechanism, and further prove that this is essentially tight. We show that this result extends to a combinatorial cost setting [57], where the cost function is an arbitrary monotone set function. We finally obtain a constant factor approximation mechanism for the general setting of DCIC for rich class of instances.

As a conceptual contribution, our work connects the Pandora's box problem and the delegated choice problem, complementing the intimate connection between the delegated choice problem and prophet inequality by [3], one of the most central problems in online algorithms. Furthermore, to the

best of our knowledge, our work initiates the first study of Pandora's box with side information given by the strategic agent's signal, which might be of independent interest.

We first briefly summarize our technical results in Section 2.2 along with a concise version of problem setup. Further related works can be found in Section 2.2.5, and formal problem setup is detailed in Section 7.2. Our main results are presented in Section 2.4, where we defer the proofs of the propositions to the appendix due to page limit.

2.2 Our results and techniques

We introduce here our main results and ideas for their proofs. To this end, we briefly introduce our problem setup. Formal model description can be found in Section 7.2. In the DCIC problem, we have a set of alternatives[1] $[n] = \{1, 2, \ldots, n\}$, and the principal needs to select one of them. Each alternative i is equipped with independent nonnegative random variables X_i and Y_i, denoting the principal's and the agent's utilities, respectively. Let D_i^X and D_i^Y denote the distribution associated with X_i and Y_i, respectively, and D^X be the product distribution of $(X_i)_{i\in[n]}$, and we similarly define D^Y. Let $\mathcal{D}^X$ and $\mathcal{D}^Y$ be the respective families of all possible such distributions over $\mathbb{R}^n_{\geq 0}$. Each alternative i is also equipped with a deterministic inspection cost c_i. The principal is aware of the distribution D_i^X and the cost c_i for $i \in [n]$. At each time, the principal can decide either to inspect an uninspected alternative, select an alternative and realize its utility, or stop the search. If the principal inspects a set of alternatives I and commits to alternative $k \in [n]$, the principal realizes the utility of $X_k - \sum_{i\in I} c_i$. Let $\mathcal{P}$ be a set of all possible search (inspection) policies.

Without the presence of the agent, the principal's search procedure to maximize the utility can be exactly framed as the Pandora's box problem with nonobligatory inspection (PNOI) [55]. On the

[1] We often refer to an alternative as a solution, an element, or a box (as in the Pandora's box problem).

other hand, the principal can consult an agent who can observe the realization of every alternative. Importantly, we assume that the principal needs to pay the exogenous cost of c_{DEL} to delegate to the agent and receive the feedback. The agent, however, realizes the utility of Y_k if the principal decides to select alternative k, *i.e.,* having a misaligned utility. Let Σ be the set of all possible signals by the agent; *e.g.,* in [3], this might be a pair of a proposed alternative and a corresponding utility. Then, given the principal's committed mechanism, the agent best-responds by sending a signal $\sigma \in \Sigma$ that maximizes the agent's utility. Overall, the game proceeds as follows:

1. The principal commits to a mechanism $\textsc{Mec} : \Sigma \to \mathcal{P}$ that maps the agent's signal to an inspection policy.

2. The agent observes the realization of $(X_i, Y_i)_{i\in[n]}$ and best-responds with signal σ given $\textsc{Mec}$.

3. The alternative (including null) is determined by the outcome of the search policy $\textsc{Mec}(\sigma)$.

We are interested in designing an *agent-oblivious* mechanism that has no information about the agent's utility distributions D^Y. Formally, the principal's expected utility for mechanism $\textsc{Mec}$ can be written as

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] = \mathbb{E}\left[\left(\sum_{i\in[n]} \mathbb{A}_i X_i - \mathbb{I}_i c_i\right) - \mathbb{I}_{\text{DEL}} c_{\text{DEL}}\right],$$

where $\mathbb{A}_i$ is the indicator random variable denoting whether i is selected, $\mathbb{I}_i$ denotes an indicator variable of whether i is inspected, and $\mathbb{I}_{\text{DEL}}$ indicates whether the principal decides to delegate or not. Furthermore, the worst-case utility of the principal for an oblivious mechanism $\textsc{Mec}$ is defined over

the worst-case possible agent's distribution, written as

$$\mathbb{E}_{D_X}\left[\textsc{Mec}\right] = \min_{D_Y \in \mathcal{D}_Y} \mathbb{E}_{D_X, D_Y}\left[\left(\sum_{i\in[n]} \mathbb{A}_i X_i - \mathbb{I}_i c_i\right) - \mathbb{I}_{\textsc{Del}} c_{\textsc{Del}}\right].$$

Given an optimal mechanism Opt that maximizes the above quantity over the space of feasible mechanisms, Mec is an α-approximation if

$$\alpha \cdot \min_{D_X \in \mathcal{D}_X} \frac{\mathbb{E}_{D_X}\left[\textsc{Mec}\right]}{\mathbb{E}_{D_X}\left[\textsc{Opt}\right]} \geq 1,$$

i.e., the worst-case ratio between the optimal oblivious mechanism and Mec.

As discussed, DCIC can be viewed as a version of PNOI problem where the side information from a strategic agent is possible at certain costs. On one hand, one can regard such information from the agent as another box that changes the distribution of the other boxes. On the other hand, the correlation structure on how the agent's signal changes the other boxes' distributions is very unclear, which makes our problem significantly harder and fundamentally different from the Pandora's box problem with correlated rewards.

Before presenting our main results, it is straightforward to see that DCIC problem is NP-hard, since the PNOI problem reduces to DCIC by setting $c_{\textsc{Del}}$ sufficiently large.

Proposition 2.2.1. *DCIC problem is NP-hard.*

2.2.1 First attempt: PNOI policy or SPMI

A natural attempt that one might try first might be to see if either of the delegation-only (without further inspection on elements not proposed by the agent) mechanism or inspection-only (without

delegation), *i.e.,* a PNOI policy, mechanism is sufficient to achieve any constant approximation.

We observe that the inspection-only mechanism cannot achieve constant approximation as follows, which necessitates the a smarter mechanism beyond the delegation.

Restatement of Proposition 2.4.1. *There exists a problem instance in which the optimal direct inspection policy cannot achieve a constant approximation ratio.*

Intuitively, one can consider n identical boxes with the same distribution supported on two values $v > 0$ and 0. For the inspection-only policy without delegation, the decision maker only needs to determine how many alternatives to inspect before it observes v due to the symmetry of the alternatives. On the other hand, one may set a simple threshold-based delegation mechanism for which the agent is required to submit a single alternative and it is accepted if and only if it induces the principal's utility at least the specified threshold. If we appropriately set the probability for v, there exists only one alternative realizing v with high probability. In this case, the inspection-only policy necessarily suffers significant amount of inspection costs, whereas the delegation mechanism, with any nonzero threshold would enjoy the utility of v with just a single inspection.

For the delegation-only mechanism, we consider the single-proposal mechanism (SPM) proposed by [3], which achieves 2-approximation to the first best outcome $\mathbb{E}\left[\max_{i\in[n]} X_i\right]$ under the assumption with truthful report. Since the original SPM does not inspect the proposed solution, which incentivizes the misreport of the utility of the proposed alternative by the agent, we consider a natural variant of SPM, dubbed *single-proposal mechanism with inspection* (SPMI), that further inspects the proposed alternative. Further, the following result shows that the single-proposal mechanism fails to achieve any constant approximation ratio as well.

Restatement of Proposition 2.4.2. *There exists a problem instance in which the SPMI cannot be*

α-approximate for any $\alpha > 0$.

The result follows from the intuition that if the inspection cost is sufficiently large, then selecting an alternative without any inspection dominates the SPMI.

2.2.2 Next attempt: using ideas of [28, 29]

On the other hand, one may wonder whether we could smartly upper bound the optimal mechanism in an efficient manner, to build our intuition towards an efficient mechanism. Indeed, such approaches are quite common in the original Pandora's box problem [10], and even for its variant [29]. These techniques, however, turn out to not easily extend to DCIC, as will be shown in what follows.

Let us first recap the standard approaches. In the standard Pandora's box problem, the principal's utility using the optimal policy can be upper bounded by a tuple of random variables $(\kappa_i)_{i\in[n]}$ defined by $\kappa_i = \min(X_i, \sigma_i)$ where σ_i satisfies

$$\mathbb{E}\left[\,(X_i - \sigma_i)^+\,\right] = c_i.$$

The quantity σ_i is usually called *cap*. Then, [28] proved that

$$\mathbb{E}\left[\,\mathbb{A}_i X_i - \mathbb{I}_i c_i\,\right] \le \mathbb{E}\left[\,\mathbb{A}_i \kappa_i\,\right]. \tag{2.1}$$

The proof follows from decomposing X_i into $\kappa_i + b_i$ such that

1. If $X_i > \sigma_i$, then *capped value* $\kappa_i = \sigma_i$ and *bonus* $b_i = X_i - \sigma_i$.

2. Otherwise, $\kappa_i = X_i$, *i.e.,* capped value is claimed.

Then, one can easily see that the construction of the bonus b_i satisfies $c_i = \mathbb{E}[\, b_i \,]$, *i.e.,* the expected amount of bonus marginalizes the cost of the inspection. Thus, a net utility from each box can be viewed as exactly the capped value.

Correspondingly, this inspires a very natural greedy-like search policy to match the described upper bound: opens the boxes in a decreasing order of σ_i, and stops when the largest observed X_i exceeds the highest remaining cap σ_j and select such i. This policy is shown to be exactly the optimal and achieving the above upper bound.

For the PNOI problem, [29] analogously follows this intuition to upper bound the optimal mechanism's utility in a structured way. Indeed, one can modify the random variable κ_i to $\tilde{\kappa}_i$ so that $\tilde{\kappa}_i = \kappa_i$ if i is inspected and $\tilde{\kappa}_i = \mathbb{E}_{v_i \sim D_i^X}[v_i]$ if i is not inspected, *i.e.,*

$$\tilde{\kappa}_i = \begin{cases} \kappa_i & \text{if } \mathbb{I}_i = 1 \\ \mathbb{E}[\, X_i \,] & \text{if } \mathbb{I}_i = 0 \end{cases}$$

and maintain the upper bound by replacing κ_i into $\tilde{\kappa}_i$, *i.e.,* $\mathbb{E}[\, \mathbb{A}_i X_i - \mathbb{I}_i c_i \,] \leq \mathbb{E}[\, \mathbb{A}_i \tilde{\kappa}_i \,]$. Intuitively, such an upper bound easily extends from the standard setting since the reward of an uninspected box can be simply upper bounded by the expected value of its reward. Given this, a simple upper bound for the utility of the optimal PNOI algorithm can be found based on the solution to Pandora's box problem and the fact that the maximum utility gained from choosing uninspected solutions, $E[\sum_i X_i \mathbb{A}_i \cdot \mathbf{1}\{\, \mathbb{I}_i = 0 \,\}]$, is at most $\max_{i \in [n]} \mathbb{E}[\, X_i \,]$.

On the other hand, in our setting, even the uninspected solutions' utility can be significantly larger than $\mathbb{E}[\, X_i \,]$ depending on the principal's mechanism and the information gained from the agent's signal, formally stated as follows.

Restatement of Proposition 2.4.3. *There exists a problem instance where the ratio of utility gained from uninspected solutions for some mechanisms and corresponding agent's best-response signal,* i.e., $E[\sum_i X_i \mathbb{A}_i \cdot \mathbf{1}\{\mathbb{I}_i = 0\}]$ *to the maximum expected utility of any solution* $\max_{i\in[n]} \mathbb{E}[X_i]$ *is* $\Omega(n)$.

Thus, it is highly unclear how to upper bound the optimal mechanism's utility for DCIC unlike the PNOI problem or the standard Pandora's box problem.

2.2.3 Costless delegation

We will now outline our approach that will lead to constant approximate mechanisms for DCIC. Recall that in previous attempts, we observe that it is crucial to (i) obtain a reasonable upper bound on the optimal mechanism, and (ii) accordingly design an efficient mechanism that smartly exploits both the delegation and further inspection. To this end, we first consider a simpler case where the cost of delegation is zero, which will serve as a stepping stone in the general setting in Section 2.2.4.

We first formally provide our main result in the costless delegation setting.

Restatement of Theorem 2.4.8. *There exists a mechanism that achieves an approximation ratio of 3 for delegated choice with inspection costs problem if the delegation cost is zero.*

Let us briefly summarize our technical steps to prove this theorem. Clearly, the principal is always not worse off by choosing to delegate and receive the agent's signal since she can simply ignore the signal in any case.

Thus, we first aim to establish an upper bound for the utility obtained from an optimal mechanism. To this end, we introduce a notion of *overinspection*. Formally, an alternative i is selected without overinspection if

$$\mathbb{A}_i = 1, \quad \mathbb{U}_i = \bigvee_{j:c_j \geq c_i} \mathbb{I}_j = 0,$$

i.e., any alternative with the same or higher inspection cost than c_i has not been inspected. Then, we can effectively prove that there exists an agent such that the expected utility from selecting a solution without overinspection is upper bounded by the expected value, *i.e.,*

$$\sum_{i\in[n]} \mathbb{E}\left[\, X_i \mathbb{A}_i \cdot \mathbf{1}\left\{\, \mathbb{U}_i = 0 \,\right\}\right] \leq \max_{i\in[n]} \mathbb{E}\left[\, X_i \,\right]. \tag{2.2}$$

At the heart of this proof is to construct a best-response from an agent such that if a solution is chosen without overinspection, it is always the lowest-cost solution that can be selected without overinspection given the principal's mechanism. Note that (2.2) exhibits a similar structure to (2.1), but has subtle differences on both sides which allow us to further upper bound the utility of any policy.

Accordingly, we provide an upper bound on the utility obtained without overinspection from this best-response by considering another response that leads to the selection of a non-overinspected solution whenever possible. Finally, we argue that whether a solution is chosen without overinspection given this response is dependent only on the utility of alternatives with strictly lower inspection costs, and therefore independent of the utility of the selected alternative itself. This independence shows that we essentially gain no information on the utility of the selected alternative and the expected utility gained will match the alternative's ex-ante utility $\mathbb{E}\left[\, X_i \,\right]$.

Combining these ideas, in Lemma 2.4.5, we can further upper bound the utility obtained by any mechanism with the following quantity:

$$\mathbb{E}\left[\, \textsc{Opt} \,\right] \leq \max_{i\in[n]} \mathbb{E}\left[\, X_i \,\right] + \mathbb{E}\left[\, \max_{i\in[n]} (X_i - c_i)^+ \,\right], \tag{2.3}$$

where $\textsc{Opt}$ is an arbitrary (optimal) mechanism's expected utility.

Given the upper bound on any feasible mechanism, it remains to derive a mechanism that is a constant approximate to the quantity in (2.3). We consider a simple form of *maximal mechanism* that takes the one with the higher expected utility from two mechanisms. Those two mechanisms are (i) $\textsc{Mec}_1$: the one that uses the SPMI, and (ii) $\textsc{Mec}_2$: the one that simply selects the solution with the largest expected utility without any inspection.

We first prove that $\textsc{Mec}_1$ achieves at least half of the last term in the right hand side of (2.3) in Lemma 2.4.6. Then, we can prove that the maximal mechanism described above achieves the maximum between the following two quantities:

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \geq \max\left(\max_{i\in[n]}\mathbb{E}\left[\,X_i\,\right], \frac{1}{2}\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right]\right)$$

Combining the results, we obtain the 3-approximate mechanism in Theorem 2.4.8. Further, this bound is shown to be tight for our mechanism, *i.e.,* the mechanism above is exactly 3-approximation, as stated in Theorem 2.4.11.

Interestingly, we prove that our analysis seamlessly carries over to the *combinatorial cost* setting, where there exists a monotone cost (set) function $c : 2^{[n]} \to \mathbb{R}_{\geq 0}$ that maps a set of selected elements to the overall cost that might not be additive over each item [57]. This is formally stated in Corollary 2.4.9.

Further, if the cost of the delegation is the same, *i.e.,* $c_i = c_j$ for $i, j \in [n]$, we improve the approximation ratio to be 2 in Proposition 2.4.10. This also holds in the combinatorial cost setting, if $c(\{i\}) = c(\{j\})$ for every $i, j \in [n]$.

2.2.4 Costly delegation

In the costly delegation setting, we prove the following theorem:

Restatement of Theorem 2.4.12. *Let* $\textsc{Mec}_P$ *be the optimal mechanism for the PNOI problem. If* $c_{\textsc{Del}} = \alpha\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]$ *for some* $\alpha < 1/2$*, then there exists a* $(3-4\alpha)/(1-2\alpha)$ *approximate mechanism. If* $c_{\textsc{Del}} \geq \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] - 2\mathbb{E}\left[\textsc{Mec}_P\right]$*, then there exists a* $2-\varepsilon$ *approximate mechanism for any* $\varepsilon > 0$*.*

This theorem essentially characterizes the constant factor approximability when the cost of delegation is relatively large or small. One immediate corollary of this theorem is that if $\mathbb{E}\left[\textsc{Mec}\right] \geq (1+\alpha)/2 \cdot \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]$ for some constant α, then there exists a constant factor approximate mechanism regardless of the delegation cost.

The assumption on the delegation cost is innocuous in a sense that DCIC still remains NP-hard under such assumptions, since the second case of Theorem 2.4.12 allows us to increasing $c_{\textsc{Del}}$ as desired. The case where the cost of delegation is moderately large remains a major open problem.

In its proof, we consider a maximal mechanism that either runs (i) PTAS for the PNOI problem by [58] and [56] or (ii) the SPMI that only delegates and inspects the (nonnull) proposed element. The proof begins with decomposing the optimal mechanism's utility similar to the costless setting, and exploits the $(1-\varepsilon)$ approximability of PTAS and the imposed assumption in a careful manner. Finally, we provide an improved 2-approximate mechanism when the cost of inspection c_i is the same for every $i \in [n]$.

2.2.5 Related Works

[1] initiates a substantial line of research on delegation mechanisms without transfers. They examined a setting where a principal delegates an optimization problem over a compact interval to a single agent and established conditions for the existence of an optimal mechanism for the principal. [59] expand on these findings by offering a broader characterization of the optimal mechanism. The discrete choice setting is first explored by [2] in which an agent samples potential solutions from a distribution and optimizes over this set, providing a setup closely aligned with our problem.

Recently, [3] identify its novel connection to the prophet inequality, and analyze several approximation factor to the first-best outcome. We note here that the second model of [3], termed delegated search problem, also reveals its connection to the Pandora's box problem. However, their model is significantly different from ours since in their model, since the agent suffers the search cost to examine the alternative assuming that the agent truthfully reports the utility afterwards, whereas in our model the principal handles the search cost and the agent also might strategically misreport the utility. [60] also tackle the assumption of truthful reports of the proposed option, or alternatively easily verifiable utility of the proposed options, with a stylized model on the principal's and the agent's utility. However, they consider a slightly different model under which the principal can only rule each option completely in or out based on the distributional knowledge, without relying on additional device such as inspection. Interestingly, they also show that the approximation to the first-best benchmark is not possible similar to ours, and thus focus on second-best benchmark. They show that the optimal mechanism is NP-hard to compute, and provide $1/3$-approximate deterministic mechanism. A number of papers [13, 47, 61, 62] study several other variants of the delegated choice problem from algorithmic perspective, but we will omit the details due to significant difference.

The Pandora's box problem, initiated by [10], has been also a central problem in theoretical computer science, grounded as one of the most fundamental online algorithm problem. The optimal policy is surprisingly simple to have an index-based nature. The most closely related to ours is the variant with nonobligatory inspection introduced by [55],[2] where the searcher can select a box without necessarily opening it. [29] propose a class of policy defined as committing policy, and prove that the optimal committing policy achieves $1 - 1/e$ approximation to the optimal policy. Simultaneous works by [58] and [56] prove that the nonobliatory inspection problem is NP-complete, and provide a PTAS algorithm based on the framework by [64]. Another line of works study several other variants of the standard Pandora's box problem [65, 66], however, their setup is largely different from ours.

2.3 Problem Setup

2.3.1 Delegated Choice

We first introduce the *delegated choice* problem by [2, 3]. There is a principal who wants to solve a certain problem, for which there is a set $\omega = \{\omega_1, \omega_2, \ldots, \omega_n\} \subseteq \Omega$ of possible pre-specified alternatives (equivalently, solutions, options, or boxes). However the principal is not an expert and may want help from an expert agent to choose the right solution. Without loss of generality, we assume Ω includes a null element $\perp$ to handle the case when no element is selected in the end, which results in the utility of zero for both the principal and agent. Each ω_i is independently sampled from a probability distribution D_i supported on Ω. We assume that the set ω of available elements (including the number of solutions) is exogenous to the mechanism designed by the principal, which is standard in literature

[2] In fact, [63] firstly study the setup of the Pandora's box with nonobligatory inspection (with discrete reward distribution) in the context of multi channel wireless networks, yet [55] firstly formalizes the problem as a variant of the Pandora's box problem.

of the delegated choice problem, *c.f.,* [2, 3]. We write Ω^* to denote the collection of finite sequences over Ω.

The principal and the agent have their own utility functions $x(\cdot), y(\cdot) : \Omega \mapsto \mathbb{R}_{\geq 0}$, respectively. These utility functions might be misaligned, and the agent is interested in maximizing his own payoff, rather than maximizing the principal's. There exists an information asymmetry between the principal and the agent such that the principal is not aware of the utilities, but the agent can observe all the utilities $x(\omega_i)$ and $y(\omega_i)$ for $i \in [n]$. After the agent observes the elements and utilities therein, he sends a signal, *e.g.,* an element, to the principal, and then the principal finally decides whether or not to accept the proposed element. Without any screening device to restrict the agent's behavior, the agent will obviously send a signal to maximize his own payoff, *i.e.,* propose an element that maximizes $y(\omega_i)$. We write X_i and Y_i to denote the random utility $x(\omega_i)$ and $y(\omega_i)$, respectively, and D_{X_i} and D_{Y_i} to denote the corresponding probability distributions of these utilities. Let $D_X = \times_{i\in[n]} D_{X_i}$ and $D_Y = \times_{i\in[n]} D_{Y_i}$, and $\mathcal{D}_X$ be the set of all possible distributions D_X over nonnegative real numbers given $[n]$, and similarly define $\mathcal{D}_Y$.

To prevent adversarial behavior of the agent, the principal devises a mechanism $\textsc{Mec} = (\Sigma, g)$ where Σ is an arbitrary set of signals and $g : \Sigma \to \Omega$ is an allocation function given the agent's signal. Namely, once the principal commits to a *signaling mechanism* MEC, the agent sends a signal as best-response, and the principal commits to an element based on the sent signal. We write X^* to denote the set of all possible finite sequences over X for any finite set X. Given a mechanism MEC, the agent's strategy is defined by a mapping $\sigma : \Omega^* \to \Sigma$. Namely, the agent observes the realization ω and corresponding utilities, and best-responds via $\sigma(\omega_1, \ldots, \omega_n)$. We often write σ^* to denote the agent's best-response.

To design an effective mechanism, the principal uses a prior information regarding utility of

the solution. Formally, we mostly focus on *agent-oblivious Bayesian* mechanisms which know the distribution of X_i but not Y_i for $i \in [n]$, and whose performances are measured with respect to the worst-case agent's distribution of Y_i's. Correspondingly, we assume that the principal cannot observe the agent's utility for each element in agent-oblivious mechanisms. We often consider *agent-aware* mechanisms which know both distributions of X_i and Y_i, for $i \in [n]$.

Essentially, [3] consider the following class of single proposal mechanisms instead of the general signaling mechanisms.

Definition 2.3.1 (Single-proposal mechanism)**.** In a *single proposal mechanism* (SPM), the principal announces an eligible set $E \in \mathbb{R}^2_{\geq 0}$, the agent proposes an element $\omega^* \in \Omega$, and the principal accepts the element if $(x(\omega^*), y(\omega^*)) \in E$, and rejects it otherwise.

In short, the agent proposes a single solution and the principal decides whether or not to accept based on the eligible set. Notably, they formally prove that it is without loss of generality to focus on single-proposal mechanisms due to a revelation-principle-style of argument. [3] show that there exists a SPM which achieves a constant approximation to the first-best outcome in hindsight, $\mathbb{E}\left[\max_{i\in[n]} X_i\right]$, by exploring a novel connection to a version of the prophet inequality [67].

Theorem 2.3.2 ([3])**.** *There exists a SPM with a single-threshold that yields the principal's expected utility of at least* $\frac{1}{2}\mathbb{E}\left[\max_{i\in[n]} X_i\right]$.

2.3.2 Delegated Choice Problem with Inspection Cost

Importantly, however, one might notice that the definition of SPM and results therein build upon a critical assumption such that if the agent proposes the solution ω_i, then he truthfully reports the corresponding utility $x(\omega_i)$ and $y(\omega_i)$ to the principal so that the principal can reliably believe the

reported utility and decide her action depending on it. For instance, if the agent knows that the principal would simply trust the utility reported by him, the agent may strategically misreport the utility of the proposed solution to induce the principal to accept the solution.[3]

Therefore, without any procedure to verify the proposed solution's value, the principal may realize arbitrarily bad utility if she simply accepts the proposal relying on the reported utility. Hence, we consider a scenario in which the principal can further *inspect* the elements to verify that the elements correspond to correct utilities. Under this assumption that the principal can verify the proposed element's utility, it is natural to allow the principal to further inspect the other elements, or just directly inspect the elements by herself without any delegation. Our model captures all these scenarios.

Now we formally define the problem of *delegated choice with inspection cost* (DCIC). As in the delegated choice problem presented above, there exist n alternatives $\omega_1, \ldots, \omega_n$, and the principal and the agent have corresponding utility functions $x, y : \Omega \to \mathbb{R}_{\geq 0}$. Analogously, we write X_i and Y_i to denote $x(\omega_i)$ and $y(\omega_i)$. Each alternative ω_i is equipped with an inspection cost $c_i \geq 0$ for $i \in [n]$. In addition, if the principal decides to delegate to the agent, it incurs an exogenous cost of $c_{\text{DEL}} \geq 0$ to the principal.

Regardless of whether the principal delegates or not, the principal needs to design an inspection policy described as follows.

Definition 2.3.3 (Inspection policy)**.** An *inspection policy* P is defined by a mapping $P : \mathcal{H} \to ([n] \times \mathcal{A}) \cup \{\text{stop}\}$ where $\mathcal{H} = ([n] \times \mathbb{R}_{\geq 0})^*$ denotes the set of all possible inspection histories and $\mathcal{A} = \{\text{select}, \text{inspect}\}$ denotes the set of possible actions on solutions, and $\{\text{stop}\}$ denotes the possible action of stopping without selecting any solution. In words, given any history $H \in \mathcal{H}$ that consists of

[3]The previous literature precludes such scenario by assuming that such a untruthful report might be easily verified, thus the agent would truthfully report the utility given the threat of future punishment or to maintain his reputation.

inspected elements and corresponding observed utilities, the inspection policy P determines the next action of whether to select an element $i \in [n]$ (possibly without inspecting it), inspect an element i, or stop without selecting any solution. Let $\mathcal{P}$ be the set of all possible inspection policies.

If the principal decides to delegate to the agent, the principal should further commit to a signaling mechanism, defined as follows.

Definition 2.3.4 (Signaling mechanism)**.** In the *signaling mechanism*, given a set of signals Σ, the principal commits to a mapping $\textsc{Sig} : \Sigma \to \mathcal{P}$ that maps the agent's signal $\sigma \in \Sigma$ to an inspection policy $P \in \mathcal{P}$.

Finally, the principal's overall mechanism can be fully characterized by $\textsc{Mec} = (\mathbb{I}_{\textsc{Del}}, P, \textsc{Sig})$ where $\mathbb{I}_{\textsc{Del}} \in \{0, 1\}$ denotes the indicator variable of whether the principal decides to delegate (1) or not (0), $P \in \mathcal{P}$ denotes the direct inspection policy that will be run if $\mathbb{I}_{\textsc{Del}} = 0$, and $\textsc{Sig}$ is the signaling mechanism $\textsc{Sig} : \Sigma \to \mathcal{P}$ that will be run if $\mathbb{I}_{\textsc{Del}} = 1$.

In particular, in DCIC problem, we consider a *modified* version of SPM as follows:

Definition 2.3.5 (Single-proposal mechanism with inspection)**.** In a single proposal mechanism with inspection (SPMI), the principal announces an eligible set $E \in \mathbb{R}^2_{\geq 0}$, and the agent proposes an element $\omega^* \in \Omega$. If $\omega^* \neq \perp$, then the principal inspects ω^* and selects it if $(x(\omega^*), y(\omega^*)) \in E$ and otherwise does not select any element. If $\omega^* = \perp$ then the principal does not inspect nor selects any.

Then, the timing of the game can be described as follows.

1. The principal commits to a mechanism $\textsc{Mec} = (\mathbb{I}_{\textsc{Del}}, P, \textsc{Sig})$
2. If $\mathbb{I}_{\textsc{Del}} = 0$, the principal runs the inspection policy P and the outcome is determined by P.

3. Otherwise, the agent best-responds according to strategy $\sigma^* : \Omega^* \to \Sigma$ which maximizes his own utility given SIG, and the principal runs the signaling mechanism $\text{SIG}(\sigma^*((\omega_1, \ldots, \omega_n)))$ to inspect the solutions and determine the outcome.

Given a principal's mechanism MEC and the agent's response σ^* if $\mathbb{I}_{\text{DEL}} = 1$, let $\mathbb{A}_i$ be the indicator variable denoting whether the element i is selected by the principal, and $\mathbb{I}_i$ be the one denoting whether the principal inspects i, for $i \in [n]$. Then, one can observe that the mechanism's expected utility given the agent's response can be written as

$$\mathbb{E}\left[\,\text{MEC}\,\right] = \mathbb{E}\left[\left(\sum_{i\in[n]} \mathbb{A}_i X_i - \mathbb{I}_i c_i\right) - \mathbb{I}_{\text{DEL}} c_{\text{DEL}}\right],$$

where we write $\mathbb{E}\left[\,\text{MEC}\,\right]$ to denote the principal's expected utility given MEC.

Recall that [3] considers a mechanism that approximates the first-best outcome in hindsight, *i.e.,* $\mathbb{E}\left[\max_{i\in[n]} X_i\right]$. In our setting, however, it is straightforward to see that it is impossible to approximate the first-best outcome in hindsight, formalized as follows.

Proposition 2.3.6 (Inapproximability against the first-best outcome)**.** *In the DCIC, there exists a problem instance such that no mechanism can obtain a utility of α times the first-best outcome in hindsight,* i.e., $\mathbb{E}\left[\max_{i\in[n]} X_i\right]$ *for any* $\alpha = \Omega(1/n)$.

Thus, we consider a second-best benchmark that competes with the optimal mechanism. Let OPT be an optimal mechanism that the principal could have run given the distributions of X_is, which maximizes expected utility given the worst-case agent. Accordingly, we will consider a polynomial time computable mechanism that approximates OPT. We write $\mathbb{E}\left[\,\text{OPT}\,\right]$ to denote the corresponding principal's expected utility with respect to the worst-case agent's utility. Recall that $D_X = \times_{i\in[n]} D_{X_i}$

and $D_Y = \times_{i\in[n]} D_{Y_i}$ denote the sets of all possible distributions for $(X_i)_{i\in[n]}$ and $(Y_i)_{i\in[n]}$, respectively. The principal's expected utility, given mechanism MEC and distribution D_X, can be written as

$$\mathbb{E}_{D_X}[\text{MEC}] = \min_{D_Y\in\mathcal{D}_Y} \mathbb{E}_{D_X,D_Y}\left[\left(\sum_{i\in[n]} \mathbb{A}_i X_i - \mathbb{I}_i c_i\right) - \mathbb{I}_{\text{DEL}} c_{\text{DEL}}\right]. \tag{2.4}$$

We focus on a set of deterministic mechanisms that deterministically set $\mathbb{I}_i$ and $\mathbb{I}_{\text{DEL}}$ to one or zero.[4] Given an optimal mechanism OPT, the mechanism MEC is an α-approximation[5] if

$$\alpha \cdot \min_{D_X\in\mathcal{D}_X} \frac{\mathbb{E}_{D_X}[\text{MEC}]}{\mathbb{E}_{D_X}[\text{OPT}]} \geq 1. \tag{2.5}$$

We finally impose a minor assumption such that if the agent is indifferent to the eventual outcome of two actions, *i.e.,* facing the same utility, then the agent behaves in favor of the principal to choose an action that induces a higher utility for the principal. This is innocuous since it is always better to make the principal satisfactory for the agent's reputation effect unless it hurts the agent's utility.

Pandora's box with nonobligatory inspection. Interestingly, we first observe that DCIC generalizes the Pandora's box problem with nonobligatory inspection (PNOI) by [55].

Definition 2.3.7 (Pandora's box with nonobligatory inspection)**.** In the *Pandora's box problem with nonobligatory inspection* (PNOI), there are boxes indexed by $1,\ldots,n$ each of which is equipped with a random reward X_i for $i\in[n]$. The searcher knows the distribution of the random variables in priori. Opening box i incurs cost c_i and reveals the realization X_i. Each box can only be selected after it

[4]Pandora's box problem and its variant typically has no advantage of randomizing the policy, but DCIC can have some advantages since the agent's best-response could be made different by randomization. Studying randomized mechanism remains a major open problem.

[5]Note that this differs from computing worst-case ratio over every possible $(D_X, D_Y) \in (\mathcal{D}_X, \mathcal{D}_Y)$. This is mainly due to the fact that the principal is oblivious to the agent's mechanism, thus the natural objective one could consider is to optimize the expected utility over the worst-case agent's distribution.

is opened, and selecting i yields the reward of X_i. Importantly, the searcher can also select the box without inspecting it. The searcher's objective is to maximize the expected payoff which is defined as the expected reward of the selected box minus the total inspection costs.

Essentially, each box can be viewed as an element where the reward from the box is the principal's utility for that solution. Then, without any delegation process, the principal can exactly solve the Pandora's box problem with nonobligatory inspection. We write *direct inspection policy* to denote the policy entirely based on the Pandora's box with nonobligatory inspection, without any delegation. Thus, we often use the word box to denote an element for DCIC, and vice versa for PNOI.

Indeed, if the cost of delegation c_{DEL} is far much larger than the maximal value of the support of the reward X_i's, then the principal would trivially set $I_{\text{DEL}} = 0$. This implies that we can reduce PNOI problem to an instance of DCIC. Thus, since PNOI is recently proven to be NP-hard by [56], it directly follows that DCIC is also NP-hard.

Proposition 2.3.8. *The decision version of DCIC is NP-hard.*

2.4 Main Results

This section presents our main results on the 3-approximate mechanism. Our results and intuitions therein are presented in a step by step manner, *i.e.,* we first present several examples to construct our intuition towards an efficient mechanism, and present our main techniques and analysis.

2.4.1 Warm-up

Naturally, the first question one may ask is whether either of the pure delegation mechanisms, such as SPMI or the pure PNOI policy, would achieve a constant approximation ratio. We here provide

a formal explanation on why such naive approaches would necessarily fail.

We start with the example stating that if the principal only considers a direct inspection mechanism, i.e., always choosing $I_{\text{DEL}} = 0$, there exists a problem instance such that the approximation factor can be arbitrarily large.

Proposition 2.4.1 (Direct inspection fails)**.** *There exists a problem instance in the independent utility setting in which the optimal direct inspection policy cannot achieve a better than $\Omega(n)$-approximation.*

Then, one might wonder whether it is possible to obtain a good approximation if instead of precluding the delegation and directly inspecting the solutions, the principal commits to SPMI that only inspects the single solution proposed by the agent.

Proposition 2.4.2 (SPMI fails)**.** *There exists a problem instance of DCIC in which the SPMI cannot be α-approximate for any $\alpha > 0$.*

2.4.2 3-approximation in costless delegation setting

Thus, it is necessary to construct a mechanism that considers the possibility of delegation as well as further inspection of the elements or selection without inspection. To simplify our argument, we start with a simple scenario in which the cost of delegation is zero, *i.e.,* $c_{\text{DEL}} = 0$. In this case, one can always set I_{DEL} to one with zero cost since it is always better to delegate to the agent to extract some useful information rather than not delegating, and any mechanism $\textsc{Mec} = (0, P, \textsc{Sig})$ with $I_{\text{DEL}} = 0$ can be simulated by choosing a mechanism $\textsc{Mec}' = (1, P, \textsc{Sig}')$ where $\textsc{Sig}'$ maps every signal to P, *i.e.,* ignore the agent's signal.

To analyze an approximation factor of a mechanism, it is essential to obtain a proper upper bound on the benchmark $\mathbb{E}[\,\textsc{Opt}\,]$. However, it is nontrivial to quantify what the optimal mechanism would

have extracted as additional information from the agent's signal. Indeed, the following example shows that the value of information could be significant for certain types of mechanisms.

Proposition 2.4.3. *There exists an instance of the problem and a corresponding mechanism where the ratio of utility gained from uninspected solutions given any agent's signal using this mechanism,* i.e., $E[\sum_i X_i \mathbb{A}_i \cdot \mathbf{1}\{\mathbb{I}_i = 0\}]$ *to the maximum expected utility of any solution* $\max_{i\in[n]} E[X_i]$ *is* $\Omega(n)$.

This ratio shows the value of information gained from delegation, as without delegation the maximum utility that can be gained from uninspected solutions is $\max_{i\in[n]} E[X_i]$.

Towards the constant approximate mechanism, we first establish an upper bound for the utility obtained from an optimal mechanism. We say a solution i is accepted without *overinspection* if it is selected without inspecting any solution of the same or higher inspection cost than c_i. Defining indicator variables $\mathbb{U}_i = \bigvee_{j:c_j \geq c_i} \mathbb{I}_j$, we can say solution i is accepted without overinspection if and only if $\mathbb{A}_i(1 - \mathbb{U}_i) = 1$.

Now, we prove the following lemma:

Lemma 2.4.4. *For any mechanism, there exists an agent that can best-respond in a way such that the principal's utility from solutions accepted without overinspection will be at most* $\max_{i\in[n]} \mathbb{E}[X_i]$, i.e.,

$$\sum_{i\in[n]} \mathbb{E}[X_i \mathbb{A}_i \cdot \mathbf{1}\{\mathbb{U}_i = 0\}] \leq \max_{i\in[n]} \mathbb{E}[X_i].$$

Proof. Assume WLOG that the solutions are numbered in order of increasing inspection costs, *i.e.,* $c_i \leq c_j$ for $i < j$, with ties broken arbitrarily. We consider an agent with deterministic utility values $y = (y_1, \ldots, y_n)$ such that $y_i > y_j$ when $i < j$.

For a given realization of the principal's utilities $x = (x_1, \ldots, x_n)$, let $\Sigma_u(x)$ be the set of signals

that would lead to a solution being accepted without overinspection, and for any signal $\sigma \in \Sigma_u(x)$ let $acc(x, \sigma)$ be the solution accepted using that signal. Now, we consider the following response by the agent given realization x. If $\Sigma_u(x)$ is non-empty, the agent chooses a signal $\sigma \in \Sigma_u(x)$ minimizing $acc(x, \sigma)$ and consequently maximizing the agent's utility. Otherwise, if $\Sigma_u(x)$ is empty, the agent selects an arbitrary signal.

We argue that the principal's utility gained from solutions accepted without overinspection is higher for this response, compared to any best-response of the agent. For any realization x, if the agent's best-response chooses a signal in $\Sigma_u(x)$ that leads to a solution being selected without overinspection, the same solution must be selected by our response, as any other solution being selected based on a signal in $\Sigma_u(x)$ has lower utility for the agent. If the best-response does not select a solution without overinspection, then the corresponding utility $\sum_{i\in[n]} X_i \mathbb{A}_i \cdot \mathbf{1}\{\mathbb{U}_i = 0\}$ is simply 0 for this realization since $\mathbb{U}_i = 1$ when $\mathbb{A}_i = 1$, and our statement holds. Therefore, we can use this response to upper bound the principal's utility given any best-response of the agent.

Now, for this response, consider any realization $x = (x_1, \ldots, x_n)$ where solution i is accepted without overinspection. We show that for any other realization $x' = (x_1, \ldots, x_{i-1}, x'_i, \ldots, x_n)$ that differs from x only in x_i, we will also accept solution i without overinspection. Clearly, the same signal can be used to select solution i without overinspection, so $\Sigma_u(x')$ is not empty, and a signal in $\Sigma_u(x')$ is used by the agent. Then, for solution i not to be selected, a solution j with $y_j > y_i$ needs to be chosen without overinspection using a signal $\sigma' \in \Sigma_u(x')$. Since $y_j > y_i$, we have $j < i$ and therefore $c_j \leq c_i$. This means that σ' should not inspect solution i for solution j to be selected without overinspection. But then, since x and x' differ only in x_i, using the same signal given x should also lead to solution j being accepted without overinspection, which is a contradiction. So, for any such x', solution i will be accepted without overinspection.

Since changing x_i can not affect whether solution i is selected without overinspection or not, we can conclude that $\mathbb{A}_i \mathbf{1}\{\mathbb{U}_i = 0\}$, which represents solution i being selected without overinspection, can be written as a function of the X_j variables except X_i. Therefore, $\mathbb{A}_i \mathbf{1}\{\mathbb{U}_i = 0\}$ is independent of X_i. We can now use this to show that a utility of at most $\max_{i\in[n]} \mathbb{E}[X_i]$ can be gained from the solutions accepted without overinspection:

$$\begin{aligned}
\sum_{i\in[n]} \mathbb{E}[X_i \mathbb{A}_i \cdot \mathbf{1}\{\mathbb{U}_i = 0\}] &= \sum_{i\in[n]} \mathbb{E}[X_i]\mathbb{E}[\mathbb{A}_i \cdot \mathbf{1}\{\mathbb{U}_i = 0\}] && \text{(Independence)} \\
&\le \sum_{i\in[n]} \mathbb{E}[X_i]\Pr[\mathbb{A}_i = 1] \\
&\le \sum_{i\in[n]} \Pr[\mathbb{A}_i = 1] \max_{j\in[n]} \mathbb{E}[X_j] && (\mathbb{E}[X_i] \le \max_{j\in[n]} \mathbb{E}[X_j]) \\
&\le \max_{i\in[n]} \mathbb{E}[X_i], && (\textstyle\sum_{i\in[n]} \Pr[\mathbb{A}_i = 1] \le 1)
\end{aligned}$$

where the first inequality follows from the fact that $\mathbf{1}\{\mathbb{U}_i = 0\} \le 1$ and $\mathbb{E}[\mathbb{A}_i] = \Pr[\mathbb{A}_i = 1]$. a

Next, we use this lemma to prove an upper bound for the utility obtained by any mechanism.

Lemma 2.4.5. *For any mechanism, there exists an agent such that the utility obtained by the principal given the agent's best-response is at most* $\max_{i\in[n]} \mathbb{E}[X_i] + \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]$. *That is,*

$$\mathbb{E}\left[\sum_{i\in[n]} X_i \mathbb{A}_i - \max_{i\in[n]} c_i \mathbb{I}_i\right] \le \max_{i\in[n]} \mathbb{E}[X_i] + \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right].$$

Proof. Let us denote the left-hand side of the inequality in the theorem as M. We can use the fact that

at most one $\mathbb{A}_i$ can be 1 to upper bound M as follows:

$$M = \mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i - \max_{i\in[n]} c_i\mathbb{I}_i\right] \leq \mathbb{E}\left[\sum_{i\in[n]}(X_i\mathbb{A}_i - \mathbb{A}_i(\max_{j\in[n]} c_j\mathbb{I}_j))\right].$$

Now, we have a sum over i and for each index i, the sum is only non-zero when $\mathbb{A}_i$ is 1 and solution i is selected. We use the notion of selection without overinspection we introduced to break the sum down into two cases:

$$\begin{aligned}
M &= \mathbb{E}\left[\sum_{i\in[n]}(X_i\mathbb{A}_i - \mathbb{A}_i(\max_{j\in[n]} c_j\mathbb{I}_j))\right] \\
&= \mathbb{E}\left[\sum_{i\in[n]}(X_i\mathbb{A}_i - \mathbb{A}_i(\max_{j\in[n]} c_j\mathbb{I}_j))(\mathbb{U}_i + (1-\mathbb{U}_i))\right] \\
&\leq \mathbb{E}\left[\sum_{i\in[n]}(X_i\mathbb{A}_i - \mathbb{A}_i(\max_{j\in[n]} c_j\mathbb{I}_j))\mathbb{U}_i\right] + \mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i\mathbf{1}\{\mathbb{U}_i = 0\}\right].
\end{aligned}$$

In the second case, we disregard the inspection costs to obtain an upper bound. Next, in the first case, since the summation term is only non-zero when $\mathbb{U}_i = 1$, we can safely replace the maximum cost of inspected solutions with c_i. So, we can say

$$\begin{aligned}
M &\leq \mathbb{E}\left[\sum_{i\in[n]}(X_i\mathbb{A}_i - \mathbb{A}_i c_i)\mathbb{U}_i\right] + \mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i\mathbf{1}\{\mathbb{U}_i = 0\}\right] \\
&\leq \mathbb{E}\left[\sum_{i\in[n]} \mathbb{A}_i(X_i - c_i)^+\right] + \mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i\mathbf{1}\{\mathbb{U}_i = 0\}\right] \\
&\leq \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] + \mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i\mathbf{1}\{\mathbb{U}_i = 0\}\right].
\end{aligned}$$

Finally, we can use Lemma 2.4.4 to replace the second term and achieve the desired upper bound:

$$M \leq \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] + \max_{i\in[n]}\mathbb{E}\left[X_i\right].$$

a

To obtain a 3-approximate mechanism, we mainly consider a maximal mechanism that computes the expected value of two mechanisms and takes the one that has the higher value. Those two mechanisms are fairly simple: (i) SPMI with an appropriate threshold, and (ii) that simply selects the one that has the largest expected utility, *i.e.,* selecting $\operatorname{argmax}_{i\in[n]}\mathbb{E}\left[X_i\right]$ without inspection.[6]

We first prove a lemma stating that the the variant of the SPMI achieves half of $\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]$.

Lemma 2.4.6. *There exists a SPMI, which inspects the proposed element and accepts it based on a single threshold τ, that achieves $\frac{1}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]$.*[7]

Proof. Given an instance of DCIC with $X = (X_1, \ldots, X_n)$, consider an instance of DCIC with $X' = (X'_1, \ldots, X'_n)$ such that the distributions are shifted to be $X'_i = (X_i - c_i)^+$, but the inspection costs are set to be $c_i = 0$ for $i \in [n]$. Since the inspection is given at free, this coincides with the setting without inspection cost by [3]. Thus, in this problem instance, by Lemma 2.3.2, we know that there exists a SPMI $\textsc{Mec}'$ with a threshold τ that achieves $\mathbb{E}\left[\textsc{Mec}'\right] \geq \frac{1}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]$.

Now, in the original instance, consider the SPMI that accepts a solution i after inspection if $X_i - c_i \geq \tau$. We claim that the principal's utility in $\textsc{Mec}$ is the same as in $\textsc{Mec}'$ for each realization of X and X' assuming that X_i and X'_i are coupled so that $X'_i = (X_i - c_i)^+$ for every $i \in [n]$. If there exists an eligible solution i such that $X_i - c_i \geq \tau$, we would also have $X'_i = X_i - c_i \geq \tau$ and

[6]The second mechanism will be replaced with PTAS policy for PNOI in Section 4.9

[7]We will disregard the computation required to compute τ as it's beyond our interest. In fact, this can be efficiently computable for discrete random variables, and can be approximately computable for continuous random variables.

the agent would propose his most preferred eligible solution i in both cases, with the same utility of $X_i - c_i = X'_i$. Otherwise, if no eligible solution exists, the agent submits the signal $\perp$ in MEC resulting in a utility of 0 for the principal since the agent's utility is 0 regardless of the signal and he behaves in favor of the principal in case of ties. In MEC$'$, the principal's utility will also be 0 as no eligible solution exists. Thus, we have $\mathbb{E}[\,\text{MEC}\,] = \mathbb{E}[\,\text{MEC}'\,] \geq \frac{1}{2}\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right]$. a

Now, the following lemma shows that the maximal mechanism described above has the desired lower bound on the principal's utility.

Lemma 2.4.7. *There exists a mechanism* MEC *that obtains a value of at least*

$$\mathbb{E}[\,\text{MEC}\,] \geq \max\left(\max_{i\in[n]}\mathbb{E}[\,X_i\,], \frac{1}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]\right).$$

Proof. Consider a mechanism that simply selects an element that maximizes $\max_{i\in[n]}\mathbb{E}[\,X_i\,]$ without delegation nor inspection.

Obviously the expected utility from this mechanism is $\max_{i\in[n]}\mathbb{E}[\,X_i\,]$ since the principal suffers no inspection cost nor delegation cost at all. Now, if $\max_{i\in[n]}\mathbb{E}[\,X_i\,]$, is larger than (or equal to) $\frac{1}{2}\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right]$, we run the mechanism above. Otherwise, we run the mechanism suggested by Lemma 2.4.6. Thus, this maximal mechanism achieves the maximum over two quantities. a

We are now ready to prove the following theorem.

Theorem 2.4.8. *There exists a mechanism that achieves an approximation ratio of 3 for the delegated choice with inspection cost if the cost of delegation is zero.*

Proof. By Lemma 2.4.5, we know that for an optimal mechanism OPT, we have

$$\mathbb{E}\left[\,\textsc{Opt}\,\right] \leq \max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right] + \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right].$$

On the other hand, by Lemma 2.4.7, we know that there exists a mechanism with

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \geq \max\left(\max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right], \frac{1}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]\right). \tag{2.6}$$

Thus, if $\frac{1}{2}\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right] \geq \max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right]$, we obtain

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \geq \frac{1}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right], \mathbb{E}\left[\,\textsc{Opt}\,\right] \leq \frac{3}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right],$$

which implies the 3-approximation ratio. Otherwise, we have

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \geq \max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right], \mathbb{E}\left[\,\textsc{Opt}\,\right] \leq 3\max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right]$$

which also implies the 3-approximation ratio.

a

Interestingly, we observe that the analysis carries over to the more complicated setting in which the cost function is rather combinatorial, not additive.

Corollary 2.4.9. *Consider a monotone combinatorial cost setting under which the cost of inspecting a set of solutions is given by a monotone set function $c : 2^{[n]} \to \mathbb{R}_{\geq 0}$. If the delegation is costless, then there exists a 3-approximate mechanism.*

Proof. Let $c_i = c(\{i\})$ be the cost of inspecting element i by itself. Lemma 2.4.4 is not dependent on the cost function used for the principal, so it can still be applied when considering costs c_i. Consequently, Lemma 2.4.5 can be used in this case, to show that

$$\mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i - \max_{i\in[n]} c_i\mathbb{I}_i\right] \leq \max_{i\in[n]} \mathbb{E}\left[X_i\right] + \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right].$$

Now, we can see that for any monotone function c, the principal's utility can be upper bounded as

$$\begin{aligned}\mathbb{E}\left[(\sum_{i\in[n]} X_i\mathbb{A}_i) - c(\{j \mid \mathbb{I}_j = 1\})\right] &\leq \mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i - \max_{\mathbb{I}_j=1} c(\{j\})\right] \\ &= \mathbb{E}\left[\sum_{i\in[n]} X_i\mathbb{A}_i - \max_{i\in[n]} c_i\mathbb{I}_i\right].\end{aligned}$$

Therefore, the bound in Lemma 2.4.5 can be used as an upper bound for the principal's utility.

On the other hand, our algorithm when delegation is costless uses at most one inspection, so its cost is unchanged compared to the setting with additive costs c_i. Therefore, the same algorithm achieves a 3-approximation in this setting too. a

Additionally, the following theorem shows an improved approximation ratio of 2 when all inspection costs are the same, whose proof follows from considering the mechanism in Theorem 2.4.12.

Proposition 2.4.10. *If $c_i = c_j$ for any $i, j \in [n]$, then there exists a 2-approximate mechanism for DCIC with costless delegation.*

Finally, the following theorem complements the above results by proving that the approximation ratio of 3 is tight for our algorithm for costless delegation.

Theorem 2.4.11. *There exists a problem instance of DCIC such that the mechanism described in Theorem 2.4.8 cannot have approximation factor better than $3 + \varepsilon$ for any $\varepsilon > 0$.*

Proof. For a small value of $\varepsilon > 0$, consider the following instance with 3 solutions: The first two solutions have identical distributions, where $X_i = 1/\varepsilon$ with probability ε and $X_i = 0$ with probability $1 - \varepsilon$. The inspection cost for these solutions is 0. The third solution has a deterministic utility of 1, but its inspection cost is also 1. One optimal algorithm in this instance inspects the first two solutions in order, selecting them if the utility is non-zero. If both solutions report a utility of zero, the third box is chosen without inspection. The expected utility of this algorithm is

$$\varepsilon \cdot \frac{1}{\varepsilon} + (1 - \varepsilon) \cdot \varepsilon \cdot \frac{1}{\varepsilon} + (1 - \varepsilon)^2 = 3 - 3\varepsilon + \varepsilon^2.$$

Now, our algorithm considers the two options of delegation or choosing the solution with maximum expected utility without inspection. For delegation, we look at the value

$$\frac{1}{2}\mathbb{E}\left[\max_{i \in [n]} X_i - c_i\right] = \frac{1}{2}\left(\varepsilon \cdot \frac{1}{\varepsilon} + (1 - \varepsilon) \cdot \varepsilon \cdot \frac{1}{\varepsilon}\right) = 1 - \frac{\varepsilon}{2}.$$

The solution with maximum expected utility has a utility of 1 as all solutions have an expected utility of 1. Since the maximum expected utility is higher than the utility bound for delegation, our algorithm will choose the solution with the maximum expected utility without inspection, achieving a utility of 1. So, the approximation ratio of our algorithm cannot be better than $3 - 3\varepsilon + \varepsilon^2$. As ε can be arbitrarily small, the approximation ratio of 3 is tight for our algorithm. a

2.4.3 Constant factor approximation for general setting

We finally present the following theorem in the costly delegation setting.

Theorem 2.4.12. *Let* $\textsc{Mec}_P$ *be the optimal mechanism for the PNOI problem (without delegation).*

1. *If* $c_{\textsc{Del}} = \alpha\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right]$ *for some* $\alpha < 1/2$*, then there exists a* $(3-4\alpha)/(1-2\alpha)$ *approximate mechanism.*

2. *If* $c_{\textsc{Del}} \geq \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] - 2\mathbb{E}\left[\textsc{Mec}_P\right]$*, then there exists a* $2-\varepsilon$ *approximate mechanism for any* $\varepsilon > 0$*.*

Proof. Consider a PTAS mechanism P_ε for the PNOI problem by [56] and [58] that $1-\varepsilon$ approximates $\textsc{Mec}_P$.

We consider the following mechanism $\textsc{Mec}$.

1. Compute $v_1 = \mathbb{E}\left[\textsc{Mec}_{P_\varepsilon}\right]$ of running P_ε without any delegation.

2. Compute $v_2 = \frac{1}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] - c_{\textsc{Del}}$ for running SPMI given by Lemma 2.4.6.

3. If $v_1 \geq v_2$, run the mechanism $\textsc{Mec}_P$, and otherwise run the above SPMI.

Due to the construction, our mechanism satisfies

$$\mathbb{E}\left[\textsc{Mec}\right] \geq \max\left(\frac{1}{2}\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] - c_{\textsc{Del}}, \mathbb{E}\left[\textsc{Mec}_{P_\varepsilon}\right]\right) \tag{2.7}$$

If the optimal mechanism $\textsc{Opt}$, does not delegate, we have that

$$(1-\varepsilon)\mathbb{E}\left[\textsc{Opt}\right] \leq \mathbb{E}\left[\textsc{Mec}_P\right] \leq \mathbb{E}\left[\textsc{Mec}\right],$$

thus our mechanism is $1-\varepsilon$ approximation.

On the other hand, if the optimal mechanism chooses to delegate, we can subtract c_{DEL} from the bound in Lemma 2.4.5 to say

$$\begin{aligned}\mathbb{E}\left[\,\textsc{Opt}\,\right] &\leq \max_{i\in[n]}\mathbb{E}\left[\,X_i\,\right] + \mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+\,\right] - c_{\text{DEL}} \\ &= \max_{i\in[n]}\mathbb{E}\left[\,X_i\,\right] + \mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+ - c_{\text{DEL}}\,\right].\end{aligned}$$

It is obvious that $v_1 \geq \max_{i\in[n]}\mathbb{E}\left[\,X_i\,\right]$.[8] Now we consider the first scenario where $c_{\text{DEL}} = \alpha\mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+\,\right]$. Then, from (2.7) we have

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \geq \max\left((\frac{1}{2}-\alpha)\mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+\,\right], \mathbb{E}\left[\,\textsc{Mec}_{P_\varepsilon}\,\right]\right).$$

We have the following two cases:

1. $\mathbb{E}\left[\,\textsc{Mec}_{P_\varepsilon}\,\right] \leq \frac{1}{2}\mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+\,\right] - c_{\text{DEL}} \leq (\frac{1}{2}-\alpha)\mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+\,\right]$.

2. $\mathbb{E}\left[\,\textsc{Mec}_{P_\varepsilon}\,\right] > \frac{1}{2}\mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+\,\right] - c_{\text{DEL}} > (\frac{1}{2}-\alpha)\mathbb{E}\left[\,\max_{i\in[n]}(X_i-c_i)^+\,\right]$.

[8] One can also consider a maximal mechanism that mixes any mechanism with closed box inspection in the beginning.

In the first case, we have

$$\begin{aligned}
\mathbb{E}\left[\,\textsc{Opt}\,\right] &\leq \max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right] + \mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+ - c_{\textsc{Del}}\,\right] \\
&\leq \mathbb{E}\left[\,\textsc{Mec}_{P_\varepsilon}\,\right] + \mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+ - c_{\textsc{Del}}\,\right] \\
&\leq (\frac{1}{2} - \alpha)\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right] + \mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+ - c_{\textsc{Del}}\,\right] \\
&\leq (\frac{3}{2} - 2\alpha)\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right] \\
&\leq \frac{\frac{3}{2} - 2\alpha}{1/2 - \alpha}\mathbb{E}\left[\,\textsc{Mec}\,\right],
\end{aligned}$$

implying that Mec is $(3 - 4\alpha)/(1 - 2\alpha)$-approximation.

Otherwise, we have

$$\begin{aligned}
\mathbb{E}\left[\,\textsc{Opt}\,\right] &\leq \max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right] + \mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+ - c_{\textsc{Del}}\,\right] \\
&\leq \mathbb{E}\left[\,\textsc{Mec}_{P_\varepsilon}\,\right] + \mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+ - c_{\textsc{Del}}\,\right] \\
&\leq \mathbb{E}\left[\,\textsc{Mec}_{P_\varepsilon}\,\right] + (1 - \alpha)\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right] \\
&\leq (1 + \frac{1 - \alpha}{1/2 - \alpha})\mathbb{E}\left[\,\textsc{Mec}_{P_\varepsilon}\,\right] \\
&\leq \frac{3/2 - 2\alpha}{1/2 - \alpha}\mathbb{E}\left[\,\textsc{Mec}\,\right],
\end{aligned}$$

implying that Mec is a $(3 - 4\alpha)/(1 - 2\alpha)$-approximation. Since in both cases Mec is a $(3 - 4\alpha)/(1 - 2\alpha)$ approximation, the mechanism is a $(3 - 4\alpha)/(1 - 2\alpha)$ approximation in the scenario where $c_{\textsc{Del}} = \alpha\mathbb{E}\left[\,\max_{i\in[n]}(X_i - c_i)^+\,\right]$

Now consider the second case where $c_{\text{DEL}} \geq \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] - \mathbb{E}\left[\text{MEC}_P\right]$. Then, we have

$$\mathbb{E}\left[\text{MEC}_{P_\varepsilon}\right] \geq (1-\varepsilon)\mathbb{E}\left[\text{MEC}_P\right] \geq (1-\varepsilon)\left(\mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+\right] - c_{\text{DEL}}\right).$$

In this case, we have

$$\begin{aligned}
\mathbb{E}\left[\text{OPT}\right] &\leq \mathbb{E}\left[\text{MEC}_{P_\varepsilon}\right] + \mathbb{E}\left[\max_{i\in[n]}(X_i - c_i)^+ - c_{\text{DEL}}\right] \\
&\leq \left(1 + \frac{1}{1-\varepsilon}\right)\mathbb{E}\left[\text{MEC}_{P_\varepsilon}\right] \\
&\leq \left(1 + \frac{1}{1-\varepsilon}\right)\mathbb{E}\left[\text{MEC}\right],
\end{aligned}$$

concluding that it is $(2-\varepsilon')$-approximation for any $\varepsilon' > 0$ by taking ε sufficiently small.

a

2.5 Remaining Proofs

2.5.1 Proof of Proposition 2.3.6

Proof of Proposition 2.3.6. Consider n ex-ante identical elements such that X_i's are i.i.d. random variables from a distribution D such that $X_i = 1/\varepsilon$ with probability ε and otherwise 0. Let each element i's inspection cost be $c_i = c = (1-\varepsilon)\cdot\mathbb{E}\left[\max_{i\in[n]} X_i\right]$. If the principal's mechanism inspects at least one element, her expected utility will be

$$\mathbb{E}\left[\text{MEC}\right] = \mathbb{E}\left[\left(\sum_{i\in[n]} \mathbb{A}_i X_i - \mathbb{I}_i c_i\right) - \mathbb{I}_{\text{DEL}} c_{\text{DEL}}\right] \leq \mathbb{E}\left[\max_{i\in[n]} X_i\right] - (1-\varepsilon)\mathbb{E}\left[\max_{i\in[n]} X_i\right] = \varepsilon\mathbb{E}\left[\max_{i\in[n]} X_i\right]$$

On the other hand, the principal may select an element without inspecting any element. In this case, the utility will be at most

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \leq \max_{i\in[n]} \mathbb{E}\left[\,X_i\,\right] = \mathbb{E}\left[\,X_i\,\right] = 1.$$

Note, however, that the first-best outcome in hindsight is $\mathbb{E}\left[\max_{i\in[n]} X_i\right] = (1-(1-\varepsilon)^n)\cdot 1/\varepsilon$. By plugging $\varepsilon = 1/n$, we obtain

$$\mathbb{E}\left[\max_{i\in[n]} X_i\right] = (1-(1-1/n)^n)\cdot n \geq \left(1-\frac{1}{e}\right) n,$$

where we use $(1-1/n)^n \leq 1/e$. On the other hand, we know that

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \leq \max\left(1, \varepsilon\mathbb{E}\left[\max_{i\in[n]} X_i\right]\right) \leq \max\left(1, \frac{1}{n}\cdot n\right) = 1,$$

which completes the proof. a

2.5.2 Proof of Proposition 2.4.1

Proof of Proposition 2.4.1. Suppose we have n identical solutions with the same utility distribution for the principal, where each solution has utility 1 with probability p and 0 otherwise. Additionally, suppose that the cost of inspecting each solution is c.

The probability that at least one solution is realized to have utility $X_i = 1$ is

$$\Pr\left[\max_{i\in[n]} X_i = 1\right] = 1-(1-p)^n.$$

Now, consider a SPMI with an eligibility threshold of $1/2 > 0$. Any rational agent will report a solution with non-zero utility if one exists since otherwise the utility will simply be zero. Therefore, the principal's expected utility for the delegation-only mechanism is

$$\mathbb{E}\,[\,\textsc{Mec}\,] \geq 1 - (1-p)^n - c\,.$$

Thus $\mathbb{E}\,[\,\textsc{Opt}\,] \geq \mathbb{E}\,[\,\textsc{Mec}\,] = 1 - (1-p)^n - c$.

On the other hand, we consider a direct inspection policy. Since all the solutions have identical utilities for the principal along with identical costs, and since there is only one non-zero possible utility value, the number of solutions inspected is enough to describe all possible inspection policies. So, for any number $0 \leq k \leq n$, we consider the expected utility of a policy that inspects up to k solutions, selecting the first one that achieves a non-zero utility. If no inspected solution's utility is non-zero and $k < n$, the policy selects an uninspected solution with an expected utility of p. Now, we can provide an upper bound for the utility of such a mechanism. If $k \geq 1$, we can say

$$\begin{aligned}
\mathbb{E}\,[\,\textsc{Mec}\,] &\leq \sum_{i=1}^{k} p(1-p)^{i-1}(1-ic) + (1-p)^k(p-kc) \\
&= \sum_{i=1}^{k} p(1-p)^{i-1} - \sum_{i=1}^{k}(1-p)^{i-1}c + p(1-p)^k \\
&= p\frac{1-(1-p)^k}{1-(1-p)} - c\frac{1-(1-p)^k}{1-(1-p)} + p(1-p)^k \\
&= (1-(1-p)^k)(1-\frac{c}{p}) + p(1-p)^k \\
&\leq (1-(1-p)^k)(1-\frac{c}{p}) + p
\end{aligned}$$

Now, let $p = 1/n$, $c = 2/n$, and assume $n \geq 6$. Then $1 - c/p = 1 - 2 \leq 0$, so we have for any inspection

only mechanism MEC

$$\mathbb{E}\left[\,\textsc{Mec}\,\right] \leq p = \frac{1}{n}.$$

On the other hand, we have

$$\mathbb{E}\left[\,\textsc{Opt}\,\right] \geq (1-(1-\frac{1}{n})^n) - \frac{2}{n} \geq 1 - 1/e - \frac{2}{n} \geq \frac{1}{6}.$$

Therefore, no mechanism utilizing only inspection can achieve an approximation ratio better than $n/6 = \Omega(n)$.

a

2.5.3 Proof of Proposition 2.4.2

Proof of Proposition 2.4.2. Consider any instance where the cost of each solution corresponds to its possible maximum utility for the principal. For example, we can assume an instance with solutions that have $X_i = 1$ with probability $1/2$ and 0 with probability $1/2$, and each solution's inspection cost is 1. It is clear that any SPMI that always inspects the nonnull solution proposed will have a utility of at most 0, as the inspection cost is no less than the utility of the solution. On the other hand, choosing a solution without inspection leads to a positive utility of $1/2$. So, any α-approximation is impossible using only this class of mechanisms. a

2.5.4 Proof of Proposition 2.4.3

Proof of Proposition 2.4.3. Consider a problem instance with n identical solutions where each X_i is 1 with probability ε and 0 otherwise. Then, the expected utility of each solution is ε. Now, consider

a signaling mechanism with $\Sigma = [n]$, with each signal corresponding to a solution. For signal i, the principal inspects all solutions except solution i, and accepts solution i if all other solutions have a utility of 0, and does not accept any solution otherwise. Then, any rational agent will use signal i if solution i is the only solution with a non-zero utility for the principal. It can be seen that assuming inspection costs are equal to 0, the utility gained from uninspected solutions in this scenario will be $n\varepsilon(1-\varepsilon)^{n-1}$. Then, the ratio of utility from uninspected solutions to the maximum expected utility of any solution can be written as $\frac{n\varepsilon(1-\varepsilon)^{n-1}}{\varepsilon} = n(1-\varepsilon)^{n-1}$ which tends to n as ε goes toward 0. a

2.5.5 Proof of Proposition 2.4.10

Proof of Proposition 2.4.10. Consider the optimal mechanism $\textsc{Opt} = (\mathbb{I}_{\textsc{Del}}, P, \textsc{Sig})$ and our mechanism from Theorem 2.4.8. If for any signal $\sigma \in \Sigma$, inspection policy $\textsc{Sig}(\sigma)$ selects a solution i without inspecting any solution, an agent preferring i to other solutions will always use this signal. Then, the principal's utility can be upper bounded by $\mathbb{E}[X_i] \leq \max_{j\in[n]} \mathbb{E}[X_j]$.

Otherwise, for any signal σ, the policy $\textsc{Sig}(\sigma)$ inspects at least one solution at cost c. Therefore, the expected utility obtained by the principal can be upper bounded by $\mathbb{E}\left[\max_{i\in[n]} X_i - c\right]$. By Lemma 2.4.7, our mechanism achieves a utility of at least $\max(\max_{i\in[n]} \mathbb{E}[X_i], \frac{1}{2}\mathbb{E}\left[\max_{i\in[n]} X_i - c\right])$, which is at least a 2-approximation to the optimal in each case. a

Chapter 3: Delegating to Multiple Agents

This chapter presents the main results of [13] (EC 2023).

3.1 Introduction

In many real-world situations, a decision-maker (called ”principal”) is encountered with a difficult task that cannot be solved by herself.[1] To cope with it, she delegates the process of finding a solution to several agents who have the ability to solve the task, and then examines agents' proposed solutions to take the best one. An interesting tension arises when the utility of the agents are not aligned with that of the principal. Since the agents are self-interested to maximize their own payoff, they may not propose the solutions for sake of the principal.

As a concrete example, consider a research proposal selection process by a grant agency. In this problem, a *principal* (*e.g.,* the grant agency) introduces a topic for proposals and asks each *agent* (*e.g.,* a principal investigator) to send a private proposal, or a set of proposals, based on a set of his research directions (possibly coming from some distribution) related to the topic introduced by the principal. Each proposal has some valuation for the agent and some valuation for the principal which can be different (*e.g.,* a proposal may be more aligned with a research direction by the agent but less aligned with the principal's interests). The goal of the principal is to *maximize* own valuation of the finally selected proposal while the goal of each agent is to maximize his utility which is his valuation for a proposal selected by the principal and zero otherwise.

In this example, the principal faces a problem of *mechanism design without money* to incentivize the agents to behave more in favor of her. In practice, the setting here can be an *incomplete information* game in the sense that though an agent knows the topic introduced by the principal and general research topics of the other agents (*i.e.,* distributions of possible proposals), he may not know the exact topics that other agents consider for their proposals. In addition, the principal may have an *examination cost*

[1] Throughout the paper we use feminine pronouns to denote the principal and masculine pronouns to denote the agents.

(*e.g.,* travel cost for time-consuming site visits) for each proposal which should be accounted for in the proposal selection process as well. A fundamental question in this example is how should the principal design a mechanism to maximize her own utility function.

Initiated by [1], there exists a rich line of literature studying the theory of *delegation without monetary transfer*. The restriction that the principal cannot commit to a contingent monetary transfer in the mechanism imparts a restrictive structure on feasible mechanisms in which the mechanism simply commits to a set of acceptable solutions, and only decide whether or not to accept the proposed solutions. [1] study the optimality of these simple mechanisms and the characterizations of the regimes for these mechanisms to be optimal. Subsequent works including [59] and [2] develop the theory of delegation in this context and broaden its scope in fruitful direction.

Recently, with a lens of computer science, [3] study a problem of designing approximate mechanisms in the *delegated search* problem.[2] They consider a simple mechanism so-called a *single-proposal* mechanism, which consists of single-round interaction between the principal and the agent. Remarkably, it achieves a 2-approximation compared to an ideal scenario in which the principal can directly select the best solution among the entire set of agents' solutions. Their analysis is based on a novel observation regarding a connection between the dynamics of the single-proposal mechanism and the dynamics of threshold-based algorithms in *prophet inequalities* (see [6], [67], [33], [39], [41], [42] for more reference to prophet inequalities).

As we discussed in the above example, there exists a variety of real-world situations, spanning economical applications to modern intelligent systems, such that the principal needs to delegate the task to multiple agents with misaligned payoff. Surprisingly, however, the problem of delegating to

[2]They study two types of model, one of which is similar to delegated choice without search cost by [2], and the other is that with search (sampling) cost. Our model can be framed as either of a delegated search problem with exogenous solutions, or delegated choice problem which delegates the choice between given alternatives.

multiple agents in the framework of mechanism design without money has received less attention in the community. The recent work by [68], preceded by [69], study the optimal delegation mechanism with two agents in the context of mechanism design without money, but their model is significantly different from ours, and also is restricted to small number of agents with specific utility functions.[3]

In this context, we introduce the *multi-agent delegation* problem, and conduct a comprehensive study on the approximation factors of various mechanisms without monetary transfer. Technically, we aim to design a multi-agent delegation mechanism that improves upon the guarantees of a single-agent mechanism by [3]. Surprisingly, different from the single-agent mechanism which assumes the prior knowledge on the agent's distribution to derive a constant approximation factor, *i.e.,* **Bayesian mechanism** (BM), we show that we can construct an efficient mechanism even without those prior knowledge, *i.e.,* by **prior-independent mechanism** (PIM), which is mainly due to the competitive tension arise from the multi-agent extension. Notably, however, we further observe that the amount of such a compelling power from competition differs with respect to the information available to the agents, and the degree of correlation between the principal's and the agent's utility. Overall, we comprehensively analyze the approximation factors of BM and PIM in both the **complete** and **incomplete information** settings (in terms of the agents' information). We note that our main technical results lie on the *analysis of approximation factors of PIM*. To highlight a few of the main results,

1. In the complete information setting, we show that there exists a PIM such that the principal's expected utility is bounded below by the *first-best solution* (in the principal's view) of the *second-best agent*. We further quantify this approximation factor by analyzing the upper and lower bound of the principal's expected utility in various regimes. This tends to asymptotically optimal in certain regimes, *i.e.,* the gains from multi-agent delegation is significant. En route to

[3] We provide the detailed discussion in Section 3.3.

this analysis, we present several technical results regarding the order statistics of i.i.d. random variables, which might be of independent interest.

2. In the incomplete information setting, we show that under symmetric agents with uniform distribution $U[0,1]$ and the assumption of independent utility, we can construct a PIM in which there exists a $(1 - e^{-n^2/2(n-1)^2})$-approximate *Bayes Nash equilibrium* such that the principal's expected utility becomes exactly optimal, where $1 - e^{-n^2/2(n-1)^2}$ converges to $1 - 1/\sqrt{e} \cong 0.3935$ as the number of agents n increases. We numerically verify that this guarantee is almost tight. Using a similar technique, we analyze that there exists a *Bayes correlated equilibrium* with a multiplicative approximation factor of $e^{n^2/2(n-1)^2} \cong \sqrt{e}$.

3. We further show that the assumptions of symmetric agents and independent utility are necessary. More formally, without these assumptions, we can construct a problem instance such that there exists a *Bayesian Nash equilibrium* in which the agents propose their worst solutions for the principal, which may induce arbitrarily bad principal's utility.

To select the best solution among the solutions from multiple agents, however, the principal suffers a cost of *examining the candidates*. In this context, we extend our problem setup and further study a trade-off between such an examination cost and an efficiency of mechanisms, and provide an analogous result in the complete information setting.

The remainder of the paper is organized as follows. In Section 3.2, we summarize overall results and the techniques therein. Section 3.3 provides detailed discussions on the related work. Section 3.4 introduces the formal problem setup. In Section 3.5 and 3.6, we provide the results in the complete and incomplete information setting, respectively. Finally in Section 3.7, we introduce the examination cost and some results in this extension.

3.2 Overview of results

Before presenting the summary of our results, we briefly discuss our problem setup and some preliminaries. We refer to Section 3.4 for a more formal problem setup. In *multi-agent delegation* problem, there exists a principal who wants to find a solution for a task by designing a mechanism that delegates the task to n agents. Each agent i observes k_i solutions by sampling from his own distributions, where both the number of solutions and the distributions are *exogenous* to the mechanism. These exogenous parameters can be interpreted as the intrinsic level of effort and skill of the agent. Afterward, each agent proposes a subset of observed solutions (signal) to the principal. Eventually, the mechanism selects a single solution, *winner*, from the proposed ones. If the solution ω is selected as the winner, the principal receives the utility of $x(\omega)$. If ω is proposed by agent i, he gets the utility of $y_i(\omega)$ while other agents realize the utility of zero. Given the mechanism, each agent i determines a (possibly mixed) strategy σ_i to maximize own utility, which maps the observed solutions and information regarding others' solutions to the signal, and plays an action based on the strategy. Importantly, we impose no restriction on the utility function $x(\cdot)$ and $y_i(\cdot)$, *i.e.,* they might be misaligned, which brings an interesting trade-off in designing the mechanism.

We introduce technical notations to summarize our results. Denote by $\bar{\omega}_i = \{\omega_{i,1}, \omega_{i,2}, \dots \omega_{i,k_i}\}$ the agent i's observed solutions for $i \in [n]$.[4] Define $X_{i,j} = x(\omega_{i,j})$ for $i \in [n]$ and $j \in [k_i]$. Sort $X_{i,j}$ for $j \in [k_i]$ in decreasing order, and denote by $X_{i,(1)} \geq X_{i,(2)} \geq \dots \geq X_{i,(k_i)}$. We denote by $X_{i,(1)}$ the first-best solution of agent i. Again, sort $X_{i,(1)}$ for $i \in [n]$ in decreasing order, and obtain $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$. Let the owner of $X_{(i)}$ be the i-th best agent. We often overwrite $X_{(1)}$ by $X_{\max}$. Our benchmark of the mechanism is $\mathbb{E}\left[X_{(1)}\right]$, *i.e.,* the case when the principal can select the

[4] We use $[n]$ to denote $\{1, 2, \dots, n\}$ for $n \in \mathbb{N}$ throughout the paper.

best one by directly accessing all the observed solutions. Let $f_{M,\sigma}$ be an interim allocation function which specifies the outcome under the mechanism M and some strategies σ, given a realization of the solutions. Then, given some strategies σ of the agents, we say that mechanism M is (ρ, γ)-approximate if

$$\rho \mathbb{E}\left[f_{M,\sigma} \right] + \gamma \geq \mathbb{E}\left[X_{(1)} \right].$$

We denote by ρ and γ the multiplicative and additive approximation factors, respectively. We say that M has multiplicative *price of anarchy* (POA_m) of ρ if it is $(\rho, 0)$-approximate under *any Nash equilibrium* of agents' strategies, and it has multiplicative *price of stability* (POS_m) of ρ if there exists *a Nash equilibrium* such that M is $(\rho, 0)$-approximate. We similarly define POA_a and POS_a to denote the price of anarchy/stability for $(1, \gamma)$-approximation, *i.e.,* additive approximation factor.[5]

3.2.1 Contributions and techniques

We here summarize our key results and the major techniques therein. Basically, we analyze two types of mechanisms: (1) Bayesian mechanism (BM) with knowledge of the prior distributions, and (2) prior-independent mechanism (PIM) which does not know the distributions in advance. We further consider complete and incomplete information settings. In the complete information setting, each agent knows the others' observed solutions while in the incomplete information setting, he only knows the distributions from which the others sample, but not the realized solutions.

We begin by showing that any mechanism can be reduced to a *multi-agent single-proposal* mechanism (MSPM, Definition 3.4.7) in Theorem 3.4.11. Essentially, this theorem implies that we can

[5] See Definition 3.4.4 and 3.4.5 for more details.

narrow down our scope of interest to the space of the MSPMs. We note that the MSPM can be either BM or PIM depending on whether the mechanism uses the knowledge on the distributions or not. All the positive results we present below will be built upon MSPMs.

General result in the complete information setting. We start with the general result on MSPM.

Restatement of Theorem 3.5.1. *In the complete information setting, for any* $\tau \geq 0$*, there exists a MSPM with* POA_a *of*

$$\mathbb{E}\left[X_{(1)} - X_{(2)} 1[X_{(2)} \geq \tau] - \tau 1[X_{(1)} \geq \tau > X_{(2)}] \right].$$

Intuitively, the first-best agent needs to guarantee that he will be selected by the principal while maximizing his own utility. This motivates him to propose a solution that yields the principal's utility at least that of the first-best solution of the second-best agent. The proof is built upon such simple intuition, however, it is still tricky to characterize arbitrary (possibly mixed) Nash equilibrium of the agents, and the formal proof can be found in Appendix B.6.2. This result implies that the additive approximation factor is characterized by the difference between $X_{(1)}$ and $X_{(2)}$, *i.e.,* the difference of first-best solution between the first-best and second-best agents. On top of it, the principal can also smartly define the threshold τ using the prior information, and further improve the quality of the candidates, although it may decrease the probability of the acceptance.

BM in the complete information setting. From the general result, we obtain the following result on BM which recovers the single-agent guarantee in [3].

Restatement of Theorem 3.5.3 and 3.5.5. *In the complete information setting, there exists a BM with* POA_m *of* 2*. In addition, this multiplicative approximation factor is tight,* i.e., *there exists a problem*

Information	Bound	General	Symmetric α-decaying
Complete	Lower	$\textsc{PoS}_m \geq 2$ (Thm 3.5.5)	open
	Upper	$\textsc{PoA}_m \leq 2$ (Thm 3.5.3)	$\textsc{PoA}_m \leq \left(\frac{\alpha nk}{\alpha nk+1}\right)^{\frac{-1}{\alpha nk}}$ (Thm 3.5.8)
Incomplete	Lower	$\textsc{PoS}_m \geq 2$ (Cor 3.6.2)	open
	Upper	$\textsc{PoA}_m \leq 2$ (Cor 3.6.2)	$\textsc{PoA}_m \leq \left(\frac{\alpha nk}{\alpha nk+1}\right)^{\frac{-1}{\alpha nk}}$ (Cor 3.6.2)

Table 3.1: Summary of results on Bayesian mechanism. *Symmetric* refers to the symmetric agents and α-*decaying* denotes the distributions such that its c.d.f. F satisfies $F(x) \leq x^\alpha$ for some $\alpha > 0$. Lower bounds for the symmetric setting remain as open problems.

instance that no BM can achieve $\textsc{PoS}_m < 2 - \varepsilon$.

The upper bound follows from the reduction from multi-agent setting to single-agent setting, where we in fact obtain stronger comparative statics such that the multi-agent delegation is always superior to the single-agent delegation. We discuss a more detailed comparison between the single-agent and multi-agent settings in Appendix B.5. The lower bound is obtained by introducing a concept of *super-agent* equipped with the standard worst-case instance in prophet inequalities. This worst-case instance implies that there is no merit in delegating the task to multiple agents with highly *asymmetric* levels of skill.

We then naturally consider a *symmetric* case (see Definition 3.5.6) when the agents are equipped with the same parameters, *i.e.,* they have similar levels of skill and effort. In this case, for a certain class of distributions, we show that the multiplicative approximation factor can be subconstant for BM, *i.e.,* converges to 1 as n (the num. of agents) or k (the num. of solutions per agent) increases.

Restatement of Theorem 3.5.8. *In the complete information setting, consider the symmetric agents with* $X_{i,j} \sim D$ *where* D *is a distribution supported on* $[0, 1]$ *and the cumulative distribution function (c.d.f.) of* $F(x)$ *satisfies* $F(x) \leq x^\alpha$ *for some* $\alpha > 0$. *Then, there exist a BM with* $\textsc{PoA}_m$ *of* $\left(\frac{\alpha nk}{\alpha nk+1}\right)^{-\frac{1}{\alpha nk}}$. *Hence, the principal's expected utility converges to* $\mathbb{E}\left[X_{\max}\right]$ *as* n *or* k *increases.*

To obtain the subconstant bound above, we use threshold-based eligible sets that are tighter than

the ones used in the prophet inequalities. We present the summary of results for BM in Table 3.1.

PIM in the complete information setting. For PIM, we observe that Theorem 3.5.1 yields an additive approximation factor of $\mathbb{E}\left[X_{(1)} - X_{(2)}\right]$ by setting $\tau = 0$ (see Corollary 3.5.9). This indeed requires no prior knowledge as it accepts any solution. On the other hand, we argue that this guarantee can be arbitrarily bad if we do not restrict our problem instances to symmetric agents.

Restatement of Theorem 3.5.10. *In the complete information setting, suppose that principal's utility function lies in* $[0, L]$ *for some* $L > 0$*. Then for any* $\varepsilon > 0$*, there exists a problem instance such that no PIM can achieve* $\textsc{PoS}_a \leq L - \varepsilon$.

The proof is based on a construction worst-case instance with a super-agent similar to the lower bound under the general setting in BM.

We thus consider the symmetric agents and derive the following result.

Restatement of Theorem 3.5.11. *In the complete information setting, consider symmetric agents with the principal's utility function lies in* $[0, L]$ *for some* $L > 0$*. Then, there exists a PIM with* $\textsc{PoA}_a$ *of* $(1 - \frac{1}{n})^{n-1}$.

Note that $(1 - \frac{1}{n})^{n-1} \leq 1/2$ for all $n \geq 2$, and decreases to $1/e$ as n increases. This constant approximation is a big improvement over the arbitrarily bad approximation in the general setting.

Next, we improve the approximation factors of PIM to be subconstant with respect to some canonical family of distributions. At the heart of this analysis, we come up with several new results (see Lemma 3.5.12, 3.5.14, 3.5.16, and 3.5.17) in order statistics[6], which might be of independent interest. Detailed discussions are presented in Section 3.5.2.

[6]Given n i.i.d. samples, r-th order statistics refer to the r-th smallest value among the i.i.d. samples.

Restatement of Theorem 3.5.15 and 3.5.18. *In the complete information setting, consider symmetric agents such that $X_{i,j} \sim D$ where D is a distribution with a probability distribution function (p.d.f.) of $f(x)$. Let $D_{\max}$ be the distribution of $X_{i,(1)}$ for any $i \in [n]$. Denote by $\mathrm{Var}(\cdot)$ the variance of a distribution. Then, the following holds.*

1. *If D has MHR (see Section 3.5.2 for definition), then there exists a PIM with $\textsc{PoA}_a$ of $\sqrt{\frac{4\mathrm{Var}(D_{\max})}{3}}$. In case of $D = U[0,1]$, we have $\sqrt{\frac{4\mathrm{Var}(D_{\max})}{3}} = \sqrt{\frac{4k}{3(k+1)^2(k+2)}} \approx \frac{1.155}{k}$.*
2. *If $f(x)$ is nondecreasing on x, then there exists a PIM with $\textsc{PoA}_a$ of $\frac{\sqrt{12\mathrm{Var}(D_{\max})}}{n+1}$. In case of $D = U[0,1]$, we have $\frac{\sqrt{12\mathrm{Var}(D_{\max})}}{n+1} = \frac{1}{n+1}\sqrt{\frac{12k}{(k+1)^2(k+2)}} \approx \frac{3.464}{k(n+1)}$.*

Incomplete information setting. In the incomplete information setting, we first discuss that the results on the BM, Theorem 3.5.3 and 3.5.5 and Theorem 3.5.8, still carry over in the incomplete information setting. In addition, without the assumption of symmetric agents, we still have an arbitrarily bad lower bound in the approximation ratio for any PIM, *i.e.,* Theorem 3.5.10 still holds. This justifies our assumption of symmetric agents in the incomplete information setting as well. Interestingly, however, assuming the symmetric agents is not sufficient to derive a positive result.

Restatement of Theorem 3.6.4. *In the incomplete information setting, under symmetric agents with $X_{i,j} \sim U[0,1]$, there is a problem instance where no PIM has $\textsc{PoA}_a < \frac{k-1}{k+1} - \sqrt{\frac{\log n}{2k+3}}$.*

Note that by making k arbitrarily large give n, RHS can be arbitrarily close to 1. The proof is based on a construction of strongly negatively correlated utility between the principal and the agents, preceded by some algebraic manipulations. Interestingly in this problem instance, each agent's utility decreases by playing for sake of the principal. Hence, each agent settles with a solution that gives the agent largest ex-post utility (and thus smallest principal's utility), even though it yields a small

Information	Bound	General	Symmetric (Sym)	Sym MHR	Sym Inc	Sym Ind
Complete	Lower	$\textsc{PoS}_a \geq L$ Thm 3.5.10	Open	Open	Open	Remark 3.2.2
	Upper	Cor 3.5.9	$\textsc{PoA}_a \leq L\left(1-\frac{1}{n}\right)^{n-1} \cong \frac{L}{e}$ Thm 3.5.11	$\textsc{PoA}_a \leq \frac{1.15}{k}$ Thm 3.5.15	$\textsc{PoA}_a \leq \frac{3.46}{k(n+1)}$ Thm 3.5.18	Remark 3.2.2
Incomplete	Lower	$\textsc{PoS}_a \geq L$ Cor 3.6.3	$\textsc{PoA}_a \geq \frac{k-1}{k+1} - \sqrt{\frac{\log n}{2k+3}}$ Thm 3.6.4	Remark 3.2.1	Remark 3.2.1	Open
	Upper	Remark 3.2.3	Remark 3.2.1	Remark 3.2.1	Remark 3.2.1	0.39-$\textsc{PoS}$ is $\textsc{Opt}$ Thm 3.6.7

Table 3.2: Summary of results on prior-independent mechanism. *MHR* refers to MHR distribution, *Inc* refers to nondecreasing p.d.f., and *Sym Ind* refers to independent utility setting with the uniform distribution $U[0,1]$. *Open* denotes that this regime remains an open problem. For general case lower bound for both settings, and symmetric case upper bound in the complete information setting, we assume that the support of the principal's utility function lies in $[0, L]$. For Sym MHR, Sym Inc, and lower bound in the incomplete information setting under Sym, we present the results for $U[0,1]$ for brevity.

probability to be selected. We present the result for uniform distribution for simplicity, but we in fact provide a more general result.

In this context, we additionally consider an *independent utility* assumption in which the agents' and principal's utilities are independent of each other (see Definition 3.6.5 for more details). Then we show that there exists a good approximate Bayes Nash equilibrium (BNE) (see Definition 3.6.6 for more details) such that the principal's expected utility becomes exactly optimal.

Restatement of Theorem 3.6.7. *In the incomplete information setting, consider symmetric agents and independent utility assumption such that $X_{i,j}, Y_{i,j} \sim U[0,1]$. Then, there is an optimal PIM with β-approximate BNE where $\beta = 1 - \exp\left(\frac{-n^2}{2(n-1)^2}\right)$. Note that β becomes $0.3935...$ as n grows.*

Different from the complete information setting, deriving and analyzing BNE is far more challenging, especially for our problem setup. One reason is that the feasible action of an agent depends on the realization of his solution. In addition, computing the expected payoff of an agent's strategy requires us to compute the expectation over every possible realization of the others' solutions. Despite these difficulties, we obtain the above result, of which the proof heavily relies on the FKG inequality

(Lemma D.3.5) and some properties on the joint density function of two order statistics, preceded by some algebraic manipulations. Using a similar but more involved technique, in Appendix B.3, we show that there exists a PIM such that under some *Bayes correlated equilibrium*, it achieves a multiplicative approximation factor of $\exp\left(\frac{n^2}{2(n-1)^2}\right) \cong 1.649$ (as n increases). We present the summary of results on PIM in Table 3.2. We remark on some points regarding the table as follows.

Remark 3.2.1. We do not study the upper bound for symmetric agents in the incomplete information setting because the lower bound in this setting can be arbitrarily bad for certain problem instances. In addition, we do not consider Sym MHR and Sym INC in the incomplete information setting since the lower bound in the symmetric setting already implies a pessimistic result.

Remark 3.2.2. Since the complete information setting inherently admits positive results only with the symmetry assumption, we do not consider the symmetric agents and independent utility scenario in the complete information setting.

Remark 3.2.3. The upper bound for the general incomplete information setting is not studied, as one can easily obtain a mechanism with an approximation ratio of ∞, *e.g.,* a mechanism that selects a solution to maximize the agent's utility.

Examination cost. Finally, we study an extension to consider the examination cost of the mechanism. The examination cost of a mechanism is defined by the number of candidates that the mechanism has to access to choose the winner. We refer to Definition 3.7.2 for more details. In this extension, we obtain the following result of a similar flavor to Theorem 3.5.1.

Restatement of Theorem 3.7.5. *Define $X_{(n+1)} = -\infty$. For any $\tau \geq 0$, there exists a mechanism with*

examination cost $2 \leq B < n$ *such that* POA_a *is at most,*

$$\mathbb{E}\left[X_{(1)}\right] - \mathbb{E}\left[X_{(2)}1\left[X_{(2)} \geq \tau > X_{(B+1)}\right] - \tau 1\left[X_{(1)} \geq \tau > X_{(2)}\right] - \sum_{e=B+1}^{n} X_{(e-B+2)}1\left[X_{(e)} \geq \tau > X_{(e+1)}\right]\right].$$

This depicts the trade-off between the examination cost and the approximation ratio. Indeed, POA increases as B decreases, and plugging $B = n$ carries over to the bound in Theorem 3.5.1.

3.3 Related work

Initiated by [1], a rich line of literature has studied the delegation mechanism without monetary transfer. [1] considers a setting in which a single agent is delegated to solve an optimization problem over a compact interval, and characterizes conditions under which an optimal solution of the principal exists. [59] fulfill the results by providing a general characterization of the solution. [2] study a discrete model in which an agent samples a set of solutions from a distribution and optimizes over the sampled set, which is the most similar problem setup to ours.

While most literature in the theory of delegation aims at characterizing regimes in which the optimal mechanism exists or the structure of it, [3] study the loss of efficiency from delegating the task into an agent with misaligned payoff. They consider two models, one of which is similar to delegated project choice of [2]. In this model, the agent samples a fixed number of solutions from a distribution and proposes some of the solutions to the principal in a strategic manner. Surprisingly, they show that a simple mechanism enjoys a 2-approximation ratio compared to the ideal scenario in which the principal can directly choose the optimal solution, based on a connection to the prophet inequalities.

Several works have been devoted to the delegation mechanism without money in the presence of multiple agents. One of the most relevant problem setting would be *legislative game* with asymmetric information by [70], [71], [72] and [73]. They consider a model in which two self-interested agents with information advantage observe *a state of nature* and independently send a private signal to the principal. The principal aims to extract truthful information on the state of nature to maximize own utility that depends on it. Recent works by [69] and [68] extend these settings by considering multiple actions for the principal. Importantly however, all of their models are significantly different from our model. In our model, the information asymmetry between the principal and the agents arises from "moral hazard" versus "adverse selection" of the previous works, as per [74].[7] In moral hazard, the principal cannot directly observe all the alternatives, but can only access a subset of them based on the signals, whereas in adverse selection, a state of nature is revealed only to the agent but affects the principal's utility as well.

Finally, there is an increasing attention in the Computer Science community regarding an algorithm design with a delegated decision. [61] considers a stochastic probing problem when the principal delegates the role of probing to a self-interested agent. They quantify the loss of delegation similar to [3] by revealing a connection between delegated stochastic probing and generalized prophet inequalities. [47] generalizes the second model of [3], namely Weitzman's box problem which deals with a costly sampling of the agent, by considering a combinatorial constraint.

[7] Notably, [74] study both effect of moral hazard and adverse selection in the delegation mechanism but with a monetary transfer. Likewise, there exists a fruitful line of literature dealing with contract-based delegation, *e.g.,* [75] and [76], but we will not discuss in detail on the delegation mechanism with money.

3.4 Problem setup

Multi-agent delegation. Consider a *principal* who designs a delegation mechanism and a set of n *agents* who perform the task and decide solutions. Let Ω be a set of potential solutions. Each agent is equipped with a number of solutions $k_i \in \mathbb{N}$, and each solution j for $j \in [k_i]$ is sampled from a probability distribution $D_{i,j}$ supported on Ω. We assume that $k_i \geq 2$ for $i \in [n]$ to avoid trivial regime. Each distribution $D_{i,j}$ may refer to the potential quality of the jth solution of agent i, and the number of solutions k_i can be interpreted as the agent's willingness-to-pay in terms of his effort, both of which are exogenous to the mechanism. We denote by $\omega_{i,j} \in \Omega$ the agent i's j-th solution for $i \in [n], j \in [k_i]$, and $\bar{\omega}_i$ the agent i's multi-set of all solutions. We further use $\bar{\omega} = \cup_{i\in[n]}\bar{\omega}_i \in \Omega^*$ to denote the multi-set of all the solutions sampled by the agents. After agent i observes his own solutions $\bar{\omega}_i$, he sends a signal to the principal, and then the principal finally adopts one of the solutions based on the agents' signals.

For notational consistency, let *solution* denote an outcome observed by an agent, *candidates* to denote the set of solutions that the mechanism can access given the agents' signals, *winner* to denote the final solution adopted by the principal. We further say our problem to be *single-agent* problem if $n = 1$, and *multi-agent* problem if $n \geq 2$.

Mechanism. A *mechanism* $M = (\Sigma, f, S)$ is specified by a set of signals $\Sigma = \Sigma_1 \times \ldots \times \Sigma_n$ of which the i-th coordinate Σ_i corresponds to the signal from agent i, a set of possible random bits S of the mechanism, and an allocation function $f : \Sigma \times S \mapsto \Omega$ which denotes the winner selected by the principal given the signals by agents and the random bits. If the mechanism is deterministic, we overwrite the domain of the allocation as $f : \Sigma \mapsto \Omega$. We mostly deal with a deterministic mechanism but introduce a general definition for completeness. For some given random bits $s \in S$, we denote by

M_s the mechanism in which its random bits are fixed to be s. Note that we use a generalized notion of signal to denote any possible interaction from the agents to the principal. For example, agent i may simply propose a subset of his observed solutions as candidates, or he may exploit a mixed strategy of proposing solutions randomly. Throughout the paper, we use Ω^* to denote a family of all finite multi-set over Ω.

Agent's strategy. Each agent i selects a certain signal based on a *strategy* $\sigma_i : \Omega^* \mapsto \Sigma_i$. More formally, given a multi-set of the k_i solutions $\bar{\omega}_i = \{\omega_{i,1}, \ldots, \omega_{i,k_i}\}$, agent i sends a signal $\sigma(\bar{\omega}_i)$ to the principal. Importantly, agent i determines the strategy based on the information regarding the others' solutions as well as his own solutions. We elaborate more in the subsequent paragraph. The principal has a utility function $x : \Omega \mapsto \mathbb{R}_{>0}$ and each agent i for $i \in [n]$ has a utility function $y_i : \Omega \mapsto \mathbb{R}_{>0}$. If solution $\omega \in \bar{\omega}$ is selected as the winner and ω belongs to agent i, then the principal and agent i realize utilities of $x(\omega)$ and $y_i(\omega)$ respectively. The other agents $j \neq i$ realize their utility to be zero. Note that we restrict the winner to be within $\bar{\omega}$ so that the principal cannot commit to a solution beyond the observed set of solutions. We further assume that Ω includes a null outcome $\perp$. If the principal selects $\perp$ as the winner, this represents that no solution is adopted by the principal and the realized utility is zero, *i.e.,* $x(\perp) = y_i(\perp) = 0$ for $i \in [n]$.

Information structure. In terms of the principal's information for agents' distributions of solutions, we consider the following two types of mechanism.

Definition 3.4.1 (Bayesian and prior-independent mechanism)**.** We say that mechanism M is **prior-independent mechanism** (PIM), if it does not have any information on the distributions from which the solutions are drawn. If it priorly knows the distribution from which the quantity $x(\cdot)$ over the solutions are drawn for each agent, we denote by **Bayesian mechanism** (BM).

Importantly, as noticed in the definition, PIM has significant informational gain over BM. It does not require any process of acquiring information on the distribution of solutions. Indeed, in practice, it is difficult or even impossible to estimate the distribution of each agent in advance. Even if it's possible, it usually entails a costly procedure to acquire such information, and thus PIM is more broadly applicable.

Regarding the principal's information on the agents' solutions, we consider an information asymmetry arises from *moral hazard* such that the principal has no information on the realized set of solutions since she delegates the search (choice) to the self-interested agents. Instead, she can only access a subset of solutions based on the signals provided by the agents.

To model the agent's information structure, we basically assume that the agents know the principal's utility function. Furthermore, we consider two cases of complete information and incomplete information based on the agent's information with respect to the others. In the **complete information** setting, all the observed solutions are revealed to each other, thereby each agent determines his own strategy after observing all the solutions. In the **incomplete information** setting, only the distributions of the solutions are publicly available to the agents but not the exact observed solutions, and thus each agent determines his own strategy based on the given knowledge of the others' distributions.

Allocation function and overall mechanism. We first define an interim allocation function which represents the winner adopted by the mechanism given the observed set of solutions.

Definition 3.4.2 (Interim allocation function)**.** $f_{M,\sigma} : \Sigma \mapsto \Omega$ is an (interim) allocation function given mechanism M and strategy σ. That is, given $\bar{\omega} = \cup_{i\in[n]}\bar{\omega}_i$, the mechanism M selects $f_{M,\sigma}(\bar{\omega})$ as the winner when each agent i plays σ_i. In case of randomized mechanism, we use $f_{M_s,\sigma}$ to denote the allocation function given the random bits s, and $F_{M,\sigma}$ to denote the set of possible winners $\cup_s f_{M_s,\sigma}$.

Overall, the mechanism proceeds as follows: (1) the principal commits to a mechanism M, (2) each agent i observes solutions $\bar{\omega}_i$, (3) each agent i strategically sends a private signal σ_i, and (4) the principal selects the winner $f_{M,\sigma}$ (or $f_{M_s,\sigma}$).

3.4.1 Further preliminaries

We here present further notations that will be broadly used throughout the paper.

Definition 3.4.3 (Nash equilibrium)**.** Given a mechanism M, a set of strategies $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ is called a *Nash equilibrium*, if σ_i yields

$$y_i(\sigma_i, \sigma_{-i}) \geq y_i(\sigma_i', \sigma_{-i}), \tag{3.1}$$

for agent i's any other strategy σ_i' given the others' strategy σ_{-i}. We call each σ_i an equilibrium strategy of agent i. We often abuse equilibria or equilibrium to denote Nash equilibrium. Whenever the mechanism is randomized, we take expectation over the utility in (3.1).

For sake of exposition, we define $X_{i,j} = x(\omega_{i,j})$ and $Y_{i,j} = y_i(\omega_{i,j})$ for each $j \in [k_i]$ and $i \in [n]$. Furthermore, we denote by $X_{i,(j)}$ the j-th highest solution in ω_i in terms of $x(\cdot)$ so that $X_{i,(1)} \geq X_{i,(2)} \geq \ldots \geq X_{i,(k_i)}$ and define $Y_{i,(j)}$ the agent i's utility for the solution that corresponds to $X_{i,(j)}$. We often overwrite $X_{i,(1)}$ by the *first-best solution* of agent i, and the agent who has $X_{(i)}$ by *i-th best agent*. Note that $X_{i,j}$ and $Y_{i,j}$ for each $j \in [k_i]$ and $i \in [n]$ are random variables such that $X_{i,j}$ and $Y_{i,j}$ are correlated with respect to $\omega_{i,j}$.

We measure the principal's utility by comparing it to a strong benchmark OPT given a realized

set of solutions $\bar{\omega}$, defined formally as follows.

$$\textsc{Opt} = \mathbb{E}\left[\max_{\omega\in\bar{\omega}} x(\omega)\right] = \mathbb{E}\left[\max_{i\in[n]}\max_{j\in[k_i]} X_{i,j}\right].$$

We often abuse $X_{\max} = X_{(1)} = \max_{i\in[n]}\max_{j\in[k_i]} X_{i,j}$. Note that $\textsc{Opt}$ denotes the principal's utility when she can access all the solutions sampled by the agents, and select the best solution by herself.

In terms of the approximation ratio of the mechanism, we consider both the additive and multiplicative approximation factors, defined as follows.

Definition 3.4.4. Mechanism M and corresponding strategies σ are (ρ, γ)-approximate if

$$\rho\mathbb{E}\left[f_{M,\sigma(\bar{\omega})}\right] + \gamma \geq \textsc{Opt}.$$

Note that the expectation in LHS is taken over the randomness in M as well as that in solutions.

Since we deal with the approximation factor of the mechanism in its equilibrium, we further define the notion of the price of anarchy and the price of stability as follows.

Definition 3.4.5 (Price of anarchy/stability)**.** Mechanism M has multiplicative *price of anarchy* ($\textsc{PoA}_m$) of ρ if it is $(\rho, 0)$-approximate under any equilibrium. It has multiplicative *price of stability* ($\textsc{PoS}_m$) of ρ if there exists an equilibrium such that M is $(\rho, 0)$-approximate. We similarly define $\textsc{PoA}_a$ and $\textsc{PoS}_a$ to denote the price of anarchy/stability for the additive approximation factor $(1, \gamma)$.

For self-containedness, we present the mechanism introduced by [3].

Definition 3.4.6 (Single-proposal mechanism)**.** In a single-proposal mechanism (SPM) under the single-agent setting, an eligible set $R \subset \Omega$ is announced by the principal, the agent proposes a solution ω, and the principal accepts it if $\omega \in R$, and rejects it otherwise.

In our multi-agent setup, we mainly investigate a variant of the SPM defined as follows.

Definition 3.4.7 (Multi-agent single-proposal mechanism)**.** In a *multi-agent single-proposal mechanism (MSPM)*, a tie-breaking rule ρ and an eligible set $R_i \subset \Omega$ are announced by the principal for each agent i. Each agent proposes up to one solution, and the principal filters a candidate from agent i if it is not contained in R_i. After filtering, the principal selects the winner by choosing the candidate that maximizes her utility, given the tie-breaking rule ρ. We say that MSPM has a homogeneous eligible set if $R_i = R$ for $i \in [n]$.[8]

Remark 3.4.8. Importantly, in terms of the principal's information, MSPM can be both BM and PIM depending on the eligible set R_i. If we specify R_i to admit any solution from Ω, then it is PIM as we do not need any kind of prior knowledge on the solution. However, if we want to specify R_i to accept only a certain type of solution as in SPM by [3], then we essentially need the prior information regarding the distribution of the solution. This is indeed true because if we commit to an eligible set that possibly rejects some solutions without knowing the distribution of the solutions, then it is obvious that we can construct a worst-case problem instance in which the mechanism essentially cannot accept any solution, resulting in an arbitrarily bad approximation factor. We refer to Claim B.6.2 in Appendix B.6.6 for a formal discussion.

Finally, we pose a (minor) assumption that if an agent is indifferent between two solutions, he selects a solution that induces a higher utility for the principal.

Assumption 3.4.9 (Pareto-optimal play)**.** *If an agent has two solutions ω and ω' such that $x(\omega) \geq x(\omega')$, $y(\omega) \geq y(\omega')$ and at least one of these inequalities are strict, then the agent does not play strategy $\sigma_i = \omega'$.*

[8] We provide further examples of reasonable mechanisms in Appendix B.2.

Remark 3.4.10. This assumption is reasonable since there exists no merit in playing such a dominated strategy ω'. Moreover in practice, the agents may also want to behave in favor of the principal as much as possible unless they need to sacrifice their own utilities, for any future interactions.

Before presenting our main results, we begin with the following observation that any mechanism can be reduced to MSPM with the proper choice of parameters.

Theorem 3.4.11. *In the multi-agent setting, for any mechanism M with an arbitrary equilibrium σ, there exists a MSPM M' such that for any equilibrium σ', for any $\omega \in F_{M,\sigma}(\bar{\omega})$ and any $\omega' \in F_{M',\sigma'}(\bar{\omega})$, we have $x(\omega) \leq x(\omega')$ for any observed solutions $\bar{\omega} \in \Omega^*$.*

The proof is provided in Appendix B.6.1. Thanks to the theorem, it suffices to narrow down our scope of interest to the set of MSPMs. Correspondingly, we focus on analyzing the approximation factor of MSPM in both the complete and incomplete information settings.

3.5 Complete information

In the complete information setting, we first present a generic result regarding MSPM with arbitrary threshold-based eligible sets.

Theorem 3.5.1. *Given any $\tau \geq 0$, let R be an eligible set with threshold τ such that $R := \{\omega : x(\omega) \geq \tau, \omega \in \Omega\}$, and ρ be an arbitrary deterministic tie-breaking rule. Then, the MSPM with ρ and a homogeneous eligible set R has* POA_a *of at most*

$$\mathbb{E}\left[X_{(1)}\right] - \mathbb{E}\left[X_{(2)} 1\left[X_{(2)} \geq \tau\right]\right] - \mathbb{E}\left[\tau 1\left[X_{(1)} \geq \tau > X_{(2)}\right]\right].$$

We present the proof in Appendix B.6.2. To interpret the bound above, suppose that we set

$\tau = 0$. Then it implies that the principal can obtain the utility of first-best solution of the second-best agent. In addition, by properly defining the eligible set, *i.e.,* τ in this case, she can further improve the potential quality of the candidates, although it might possibly decrease the probability that each agent samples eligible solutions. Intuitively, the existence of the other agents make each agent more competitive in sending signals, thus motivating them to submit more qualified solution.

Remark 3.5.2. The role of deterministic tie-breaking rule is essential. Indeed if the principal exploits a randomized tie-breaking rule, it may reduce the power of competition in certain cases since the agents may settle with a bad solution in terms of the principal, as it still gives the agents higher utility with some small but still positive probability due to the randomized tie-breaking rule.

As per Remark 3.4.8, the mechanism presented in Theorem 3.5.1 can be either BM or PIM. In what follows, we obtain guarantees for BM and PIM independently by quantifying the approximation factor in Theorem 3.5.1.

3.5.1 Bayesian mechanism

For BM, we first show that we can exactly recover the guarantees under the single-agent setting provided in [3] by setting proper eligible sets.

Theorem 3.5.3. *There exists a BM with* $\textsc{PoA}_m \leq 2$.

The proof can be found in Appendix B.6.3. To obtain the result above, we reduce the multi-agent delegation problem to a single-agent one and then use the results by [3] regarding the upper bound on the multiplicative approximation factor.

Remark 3.5.4. In the proof of Theorem 3.5.3, we in fact derive a stronger result such that there always is a multi-agent mechanism which yields higher (or equal) principal's utility compared to any mechanism

in the corresponding single-agent setting. We complement this result in Appendix B.5 by proving that for any single-agent problem instance, we can construct a multi-agent instance with higher (or equal) principal's utility.

The tightness of the bound in Theorem 3.5.3 is also obtained as follows.

Theorem 3.5.5. *For any $\varepsilon > 0$, there exists a problem instance such that no BM achieves* $\textsc{PoS}_m < 2 - \varepsilon$.

To construct a matching lower bound, we introduce a *super-agent* of superior solutions. The super-agent always induce solutions each of which is dominant over any solution by other agents. The principal thus always select one of the super-agent's solution, which in turn reduces the multi-agent setting to the single-agent setting. Then we borrow the standard worst-case problem instance of the prophet inequalities to construct a matching lower bound. The formal proof is provided in Appendix B.6.4. In words, if the principal faces a superior agent who overwhelms all the other agents, there is no merit in delegating to multiple agents, other than the superior one.

This naturally bring us to consider a symmetric assumption defined as follows.

Definition 3.5.6 (Symmetric agents)**.** If $k_i = k$ for $i \in [n]$, and $D_{i,j} = D$ for $i \in [n], j \in [k]$ we say that the agents are symmetric.

This assumption implies that the agents have a similar level of effort and skill, and importantly, all the agents can possibly become the owner of the winner.

Example 3.5.7. In the example of research proposal selection process we presented in Section 5.1, the grant agency (principal) may announce a shortlist of the investigators (agents) with a similar level of research abilities, *e.g.,* by comparing their publication records. Indeed, the grant agency will not include an investigator who is expected to be not qualified than the others in the shortlist.

The symmetry assumption effectively sparks a fire in the competitive tension between the agents due to their similar level of quality. Formally, in the following theorem, we obtain an upper bound on the additive approximation factor of a Bayesian mechanism under the symmetric agents. We refer to Appendix B.6.5 for the proof.

Theorem 3.5.8. *Consider symmetric agents such that $X_{i,j} \sim D$ where D is supported on $[0,1]$ and its c.d.f. $F(x)$ satisfies $F(x) \leq x^{\alpha}$ for any $x \in [0,1]$ for some $\alpha > 0$. Then, there exists a BM with $\textsc{PoA}_m \leq \left(\frac{\alpha nk}{\alpha nk+1}\right)^{-1/\alpha nk}$, where $\left(\frac{\alpha nk}{\alpha nk+1}\right)^{-1/\alpha nk}$ converges to 1 as αnk increases.*

In order to obtain the above result, we introduce a tighter threshold τ compared to the ones studied in the canonical prophet inequalities algorithm. Combined with the guarantee of MSPM in Theorem 3.5.1, followed by some elementary calculations, we derive the subconstant bound above, where the principal attains the optimal utility as n and k increase. This is in stark contrast to the general setting under which the principal can obtain only a constant factor of the optimal utility.

3.5.2 Prior-independent mechanism

Now we turn our attention to the class of PIM. [9] By plugging $\tau = 0$ in Theorem 3.5.1, we directly obtain the following corollary, where we omit the proof as it is trivial to derive.

Corollary 3.5.9. *There exists a PIM with $\textsc{PoA}_a$ of at most $\mathbb{E}\left[X_{(1)}\right] - \mathbb{E}\left[X_{(2)}\right]$.*

This corollary mainly implies that even without the eligible set, the multi-agent setting brings some advantages for the principal, which is mainly due to the competitive tension between the agents. As the readers may have noticed, this additive approximation factor is equivalent to the expected dif-

[9] Standard approach to analyze prior-independent mechanism would be to use Bulow-Klemperer-style result by [16], but this does not apply for our case. We discuss more in details in Appendix B.4.

ference between the n-th (highest) order statistics and $n-1$-th order statistics of the random variables $\left(X_{i,(1)}\right)_{i\in[n]}$, which is again the highest order statistics of each agent.

Although there exists a vast amount of papers studying order statistics of similar flavor in the literature, the expected gap $\mathbb{E}\left[X_{(1)}-X_{(2)}\right]$, *i.e.,* the expected difference of the highest order statistics between the first-best and second-best agents, has not been studied before to the extent of our knowledge. In this context, we provide some analysis under which the additive factor is nicely bounded, using some characterizations on the order statistics.

Before presenting our positive results, we present the following negative result, which dictates that one should consider a reasonable assumption on the distribution to obtain positive results.

Theorem 3.5.10. *Suppose that both the principal's utility functions are supported on* $[0,L]$ *for some* $L>0$*. For any* $\varepsilon>0$*, there exists a problem instance such that* $\textsc{PoS}_a \geq L-\varepsilon$*.*

We provide the proof in Appendix B.6.6. Similar to the case of BM, we consider the symmetric agents. In this case, we can derive the following upper bound on the additive approximation factor regardless of the underlying distributions. We present the proof in Appendix B.6.7.

Theorem 3.5.11. *Under the symmetric agents, suppose that the principal's utility functions are supported on* $[0,L]$*. Then there exists a PIM with* $\textsc{PoS}_a$ *of* $L(1-1/n)^{n-1}$*, where* $(1-1/n)^{n-1}\leq 1/2$ *for all* $n\geq 2$*, and it decreases to* $1/e$ *as* n *increases.*

We also note that this upper bound $L(1-1/n)^{n-1}$ is tight for $\mathbb{E}\left[X_{(1)}-X_{(2)}\right]$, *i.e.,* we can construct a problem instance with exactly having $\mathbb{E}\left[X_{(1)}-X_{(2)}\right]=L(1-1/n)^{n-1}$. Note that this approximation ratio is a big improvement over the general setting under which the additive approximation factor can be arbitrarily bad.

Moreover, if we consider a specific scenario where the principal's utility is endowed with a certain class of distributions, we show that the approximation factor can become arbitrarily small as the number of solutions per agent (k) increases. At the heart of this result, we prove several new properties on the order statistics, which might be of independent interest. First, we show that the expected difference between any top k-th and $k+1$-th order statistics is a nonincreasing function on n for any distributions with *monotone hazard rates* (MHR). Note that a distribution D with p.d.f. $f(x)$ and c.d.f. $F(x)$ has a monotone hazard rate if $h(x) = f(x)/(1-F(x))$ is nondecreasing on x. Using a similar technique, we also show that the expected difference between any k-th and $k+1$-th lowest order statistics is also a nonincreasing function on n, for any distribution with *monotone reverse hazard rate*, where a distribution has MRHR if $h'(x) = f(x)/F(x)$ is nonincreasing on x .

In case of the difference between the first and the second order statistics, or the $n-1$-th and the n-th order statistics under the distribution with MHR, [77] shows that they indeed are nonincreasing on n with applications to the expected rent of the (reverse) auction. To the best of our knowledge, we first provide a generalized and strengthened result of theirs. Our proof builds upon the techniques by [78] which proves the convexity of k-th order statistics with respect to n. We present the precise statement below, and its proof is provided in Appendix B.6.8.

Lemma 3.5.12. *Let $X_1, \ldots, X_n$ be i.i.d. random variables equipped with a continuous distribution D. Let $X_{1:n} \leq \ldots \leq X_{n:n}$ be corresponding order statistics. Define $\Delta_{k,n} = \mathbb{E}\left[X_{n-k+1:n} - X_{n-k:n}\right]$,* i.e., *the difference between the k-th largest and $k+1$-th largest order statistics, and $\delta_{k,n} = \mathbb{E}\left[X_{k+1:n} - X_{k:n}\right]$. Then, for any $k \leq n-1$, the following holds.*

1. *If D has MHR, then $\Delta_{k,n}$ is a nonincreasing function on n.*

2. *If D has MRHR, then $\delta_{k,n}$ is a nonincreasing function on n.*

Remark 3.5.13. The fact that these quantities are nonincreasing on n is fairly intuitive at first glance since as more samples are drawn their order statistics will lie more compactly within a fixed interval, however, this is not true in hindsight for certain cases. For example, in case of Bernoulli distribution with $p = 1/2$, it is straightforward to check that $\Delta_{k,n} = \Pr\left[\sum_{i=1}^{n} X_i = n - (n-k+1) + 1 = k\right] = \binom{n}{k} 2^{-n}$. Hence, $\Delta_{k,n} \geq \Delta_{k,n+1}$ is equivalent to $\frac{n+1}{n-k+1} \leq 2$, which holds only for $k \leq \frac{n+1}{2}$.

Next, we prove that any k-th order statistics of i.i.d. random variables with MHR distribution has MHR distribution, which also has not been studied before to the extent of our knowledge. We defer the proof in Appendix B.6.9.

Lemma 3.5.14. *Let $X_1, \ldots, X_n$ be i.i.d. random variables equipped with a MHR distribution, then its r-th order statistics has MHR distribution for any $r \in [n]$.*

Using the two lemmas described above, we prove that under MHR distribution, PIM can achieve subconstant additive approximation factor as follows. We present the proof in Appendix B.6.10.

Theorem 3.5.15. *Under symmetric agents with $X_{i,j} \sim D$ such that D has MHR, there exists a PIM with* $\textsc{PoA}_a \leq \sqrt{\frac{4}{3}\mathrm{Var}(D_{\max})}$ *where $D_{\max}$ is the distribution of $X_{i,(1)}$ and $\mathrm{Var}(\cdot)$ denotes the variance of the distribution. In case of $D = U[0,1]$, the approximation factor becomes approximately $\frac{1.155}{k}$.*

We note that for uniform distribution, the subconstant approximation factor above guarantees the negative dependency on the number of solutions k. Thus, it improves largely upon the constant factor approximation of Theorem 3.5.11, and the constant factor approximation ratio of 2 for the Bayesian mechanism in [3].

Furthermore, if we consider a more restrictive class of problem instances, we prove that one can obtain both negative dependencies on n and k, *i.e.,* the PoA converges to 1 whenever n or k increases. Similar to the MHR distribution case, our proof heavily relies on the following lemmas.

Lemma 3.5.16. *Define variables the same as Lemma 3.5.12. Then, for any $k \leq n-1$, the following holds.*

1. *If $f(x)$ is nondecreasing on x, then $(n+1)\Delta_{k,n} \leq n\Delta_{k,n-1}$.*

2. *If $f(x)$ is nonincreasing on x, then $(n+1)\delta_{k,n} \leq n\delta_{k,n-1}$.*

Lemma 3.5.17. *Let $X_1, \ldots, X_n$ be i.i.d. random variables with p.d.f. f. Then, the following holds.*

1. *If $f(x)$ is nondecreasing, then the p.d.f. of the largest order statistics is nondecreasing on x.*

2. *If $f(x)$ is nonincreasing, then the p.d.f. of the smallest order statistics is nonincreasing on x.*

We highlight that Lemma 3.5.16 implies that the expected difference between two consecutive order statistics is strictly decreasing over n with the factor of $n/n+1$. Using the above lemmas, we obtain the following theorem, where the proof is presented in Appendix B.6.13.

Theorem 3.5.18. *Under symmetric agents with $X_{i,j} \sim D$ such that D has nondecreasing p.d.f. $f(x)$, there exists a PIM with $\text{PoA}_a \leq \frac{1}{n+1}\sqrt{12\text{Var}(D_{\max})}$. In case of $D = U[0,1]$, the approximation factor becomes approximately $\frac{3.464}{k(n+1)}$.*

This strictly improves upon the approximation factor of Theorem 3.5.15, since the approximation factor has positive effects on both k and n. This highlights that delegating to multiple agents indeed brings significant advantages over the single-agent case, under which the best one may hope for is to construct constant factor approximation ratio even with Bayesian mechanism. Notably, the approximation factor decreases to zero as n or k increases, *i.e.,* the principal's utility converges to $\mathbb{E}[X_{\max}]$. It remains a major open problem to see whether the guarantees above are tight or not.

3.6 Incomplete information

Recall that in the incomplete information setting, each agent i decides his own action at the interim stage of the Bayesian game, *i.e.,* when he observes his own solutions. More formally, agent i's strategy $\sigma_i : \Omega^* \mapsto \Sigma_i$ is a function from the observed solutions to a signal, which is typically a subset of solutions. We denote by $\sigma_i(\bar{\omega}_i)$ the agent's action given the observed solution $\bar{\omega}_i$. Given $\bar{\omega}_i$ and others' strategies σ_{-i}, agent i's strategy $\sigma_i(\bar{\omega}_i)$ induces the following expected utility.

$$u_i(\sigma_i(\bar{\omega}_i), \sigma_{-i}) = \mathbb{E}\left[\, y_i(f_{M,(\sigma_i(\bar{\omega}_i),\sigma_{-i})})\,\right] = \int_{\bar{\omega}_{-i}} \mathbb{E}\left[\, y_i(f_{M,(\sigma_i(\bar{\omega}_i),\sigma_{-i}(\bar{\omega}_{-i}))})\,\right] f_{-i}(\bar{\omega}_{-i}) d\bar{\omega}_{-i},$$

where $f_{-i}(\bar{\omega}_{-i})$ denotes the p.d.f. of observing $\bar{\omega}_{-i}$. Namely, each agent accounts for the probability that certain solutions are observed by the other agents, confirms their corresponding strategies, and computes his own utility based on them. This process is captured by the integral in the last equation. Note that the expectation inside the integral accounts for the random bits of the mechanism, which does not appear in deterministic mechanisms.

Then, given the others' strategies σ_{-i}, agent i's expected utility of playing σ_i can be written as

$$u_i(\sigma_i, \sigma_{-i}) = \mathbb{E}_{\bar{\omega}_i}\left[u_i(\sigma_i(\bar{\omega}_i), \sigma_{-i})\right] = \int_{\bar{\omega}_i} u_i(\sigma_i(\bar{\omega}_i), \sigma_{-i}) f_i(\bar{\omega}_i) d\bar{\omega}_i.$$

Finally, a Bayesian Nash equilibrium can be defined as follows.

Definition 3.6.1 (Bayesian Nash Equilibrium)**.** Under the incomplete information setting, given a mechanism M, a set of strategies $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is defined to be a Bayesian Nash equilibrium (BNE),

if for any $i \in [n]$, σ_i satisfies the following for any other strategy σ_i' of agent i.

$$u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i}).$$

Our main objective is to construct a mechanism that induces a large principal's expected utility under its BNE. In the case of incomplete information, the analysis of the approximation ratio becomes more involved as all the computations of the expected payoff of some strategies entail probabilistic arguments regarding the realization of the others' solutions. Moreover, since the *feasible* strategy depends on the realization of the solutions for each agent, its analysis significantly differs from the standard Bayesian games such as the first-price auction or the all-pay auction. Due to this fact, even under the symmetric agents, the standard technique of computing symmetric BNE in the first-price auction, *e.g.,* which involves the use of the Envelope theorem [79], does not apply to our setting. Despite these analytical challenges, we succeed in obtaining several positive results in terms of the PIM.

Before presenting our main result on PIM, we first study the class of the Bayesian mechanism. In the case of BM, it is straightforward to see that all the results in the complete information setting hold in the incomplete information setting as well.

Corollary 3.6.2. *Theorem 3.5.3, 3.5.5, and 3.5.8 still holds for the incomplete information setting.*

We skip the proofs as exactly the same arguments of the corresponding proofs in the complete information setting directly carry over to the incomplete information setting.

In the case of PIM, similar to the complete information setting, we can directly argue that the additive approximation factor can be arbitrarily bad. We omit the proof as it is exactly the same as the proof of Theorem 3.5.10.

Corollary 3.6.3. *Theorem 3.5.10 still holds for the incomplete information setting.*

This again naturally leads us to focus on the symmetric agents. Interestingly however, in the following theorem, we show that even under the symmetric agents, for any PIM, there exists a problem instance such that proposing a solution that minimizes the principal's utility constructs a BNE, and thus we can construct a problem instance such that the approximation factor can be arbitrarily bad. We defer the proof in Appendix B.6.14.

Theorem 3.6.4. *Under the symmetric agents, there exists no PIM such that* $\textsc{PoA}_a < \mathbb{E}\left[X_{(1)} - \max_i X_{i,(k)} \right]$. *For the uniform distribution* $x(D_{i,j}) \sim U[0,1]$, *there exists no PIM such that* $\textsc{PoA}_a < \frac{k-1}{k+1} - \sqrt{\frac{\log n}{2k+3}}$, i.e., *the additive approximation factor of any PIM can be arbitrarily bad as* k *increases given* n.

Compared to the subconstant approximation ratio of Theorem 3.5.15 and 3.5.18 in the complete information setting, Theorem 3.6.4 implies rather pessimistic consequence such that its worst-case guarantees may become arbitrarily bad, alike to the efficiency of prior-independent mechanism in the single-agent setting. Thus, the gain from multi-agent delegation becomes highly restrictive.

This negative result mainly arises from the fact that each agent has a utility function which is negatively correlated with the principal's utility, and they have limited information regarding each other's solutions. This induces the agents to propose more in a selfish manner, compared to the complete information setting.

This phenomenon resembles the *bid shading* effect in the BNEs of the first-price auction, in which the bidders tend to bid less than their valuations in equilibriums. Due to the lack of information on the other's strategy, each agent may tend to submit a candidate which brings more ex-post utility even though it has a smaller chance of being selected. Similarly from the first-price auction, this is mainly to prevent the case in which he submits more in favor of the principal when the others' solutions

were revealed to be less competitive in terms of being selected. Combined with this effect and the fact that the agent's utility is strongly negatively correlated with the principal's utility, each agent tends to submit a solution that minimizes the principal's utility.

In this context, one natural question is whether we can recover a better approximation ratio in more restricted problem instances. To this end, we consider the following assumption.

Definition 3.6.5 (Independent utility)**.** A problem instance satisfies independence if D_i is a product measure over $D_{i,x}$ and D_{i,y_i}, *i.e.,* $D_i = D_{i,x} \times D_{i,y_i}$, such that each of $x(\cdot)$ and $y_i(\cdot)$ are associated with the distributions $D_{i,x}$ and D_{i,y_i} respectively, and D_i is independent for $i \in [n]$.

Note that with independent utility assumption, we can overcome the bad problem instance presented in the proof of Theorem 3.6.4, as the principal and the agents do not longer have negatively correlated utility, but are rather not correlated at all. Although we only obtain a positive result under the independent utility assumption due to the analytical tractability, we hope that one may extend our analysis into the case where the principal and the agents have positively correlated utility (or mild amount of negative correlation), and we leave it as a major open problem.

Due to the intrinsic difficulties in analyzing BNE, we consider a more general solution concept of approximate equilibrium defined as follows.

Definition 3.6.6 (Approximate BNE)**.** Given a mechanism M, a set of strategies $\sigma = (\sigma_1, \ldots, \sigma_n)$ is ε-approximate BNE if for any $i \in [n]$, σ_i satisfies the following for any other strategy σ_i' of agent i.

$$u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i}) - \varepsilon,$$

Our definition of PoS and PoA naturally carries over to the approximate equilibrium. We even-

tually present our main result in the incomplete information setting. Our main result in the incomplete information setting is as follows.

Theorem 3.6.7. *Under the symmetric agents and independent utility with* $X_{i,j} = X \sim U[0,1]$*, there exists a PIM such that there exists* ε*-approximate BNE that results in the optimal utility,* i.e., $\mathbb{E}[X_{\max}]$*, where* $\varepsilon = 1 - e^{-\frac{n^2}{2(n-1)^2}}$*. Note that* $1 - e^{-\frac{n^2}{2(n-1)^2}}$ *converges to* $1 - 1/\sqrt{e} \cong 0.3935$ *as* n *increases.*

The proof can be found in Appendix B.6.15. Remarkably, this implies that the principal can recover the optimal utility if the agents agree to sacrifice only a constant portion of their utility. This is indeed possible due to the independence assumption and the fact that the agent does not have the exact information on the others' solutions.

Our proof is based on analyzing the probability of the event that the principal's and the agent's utility is somehow aligned, *i.e.,* an agent maximizes his expected utility by proposing his best solution in terms of principal. We show that this probability is lower bounded by $\exp\left(\frac{-n^2}{2(n-1)^2}\right)$,[10] which converges to $1/\sqrt{e}$ as n increases. The proof heavily relies on the FKG inequality (Lemma D.3.5) and some properties on the joint density function of two order statistics.

Remark 3.6.8. In appendix B.3, we also prove that using a similar but more involved technique, we can derive that there exists a PIM such that under some *Bayes correlated equilibrium*, it achieves $\left(\exp\left(\frac{n^2}{2(n-1)^2}\right), 0\right)$-approximation, *i.e.,* constant approximation ratio for sufficiently large n.

Remark 3.6.9. Note that in some cases it might be complicated or even computationally intractable for each agent to compute a best-response given the others' strategies since he needs to account for every possible realization of the solutions for the others. In this context, it would be reasonable for the agents to follow the mechanism's recommendation to propose a solution that maximizes the principal's

[10] In Appendix B.1, we numerically verify that our lower bound is almost tight.

utility, even though it may slightly decrease their utility up to some constant factors.

3.7 Examination cost

Since we allow multiple agents to propose a solution, on one hand, it helps the principal find a better solution because she can compare more solutions as observed in Section 3.5 and 3.6. On the other hand, since the principal needs to select which ones to choose among the multiple candidates, it entails an additional burden to *evaluate* the candidates, compared to the single-agent setting.

Example 3.7.1. Suppose that a project manager delegates the task of finding a solution for a project to multiple teams. Each team proposes a candidate and then the manager needs to find the best candidate by sequentially examining the candidates. This procedure can be costly for the manager and even cannot be parallelized since the manager may want to examine the candidates by herself.

To effectively study the trade-off between such a burden in evaluating the candidates and the principal's utility, we introduce the following notion of the examination cost of a mechanism.

Definition 3.7.2 (Examination)**.** Given a set of solutions $\bar{\omega} \in \Omega^*$ and corresponding strategies σ, suppose that mechanism M can eventually choose a winner among a subset $p \subset \bar{\omega}$. The mechanism cannot directly access the subset p, and does not have information on their values, but only knows from whom each candidate in p is proposed. In order to select the winner, the mechanism should examine the solution by accessing each candidate in p, and this incurs an *examination cost* of 1. We call this procedure an examination process.

It is important to note that the principal may randomize the examination process. We further denote that a mechanism has *(examination) budget* B, if its total cost of examination does not exceed B for any realized set of solutions and any corresponding strategies.

We aim to study the trade-off between the examination budget B and the approximation factor of a mechanism. Unlike the previous results in which the MSPM can examine all the candidates proposed by the agents and commit to the best one, if the examination budget is finite, the principal should selectively examine them. In this context, we present the following variant of MSPM.

Definition 3.7.3 (Randomized single-proposal mechanism)**.** In a *randomized single-proposal mechanism (RSPM)* with budget B, a tie-breaking rule ρ, an eligible set $R_i \subset \Omega$ are announced by the principal for each agent i. Each agent proposes up to one solution, and the principal selects B candidates uniformly at random and adopts the one of them to maximize her utility.

Note that RSPM with budget B always examine at most B candidates. As it can be observed in the definition of the RSPM, once a candidate $\omega \notin R_i$ is submitted by agent i, there is no chance of getting accepted by the mechanism. This implies that there is no incentive for the agent to submit a candidate beyond the eligible set. In this context, we impose the following mild assumption.

Assumption 3.7.4 (Rationality)**.** *Given a RSPM M with arbitrary tie-breaking rule ρ and eligible sets $R_1, \ldots, R_n$, each agent i for $i \in [n]$ does not play strategy $\sigma_i = \omega$ if $\omega \notin R_i$.*

In Section 3.5 and 3.6, we assume that the principal can examine all candidates (at most n) and takes the best one as the winner, *i.e.,* the principal has a sufficient budget $B \geq n$. To study the nontrivial effect of examination cost, we assume that the budget is *not sufficient* to examine all the candidates, *i.e.,* $B < n$. We analyze RPSM in the complete information setting by generalizing Theorem 3.5.1 in Section 3.5 to the limited budget case.

Theorem 3.7.5. *Given any $\tau \geq 0$, let R be an eligible set with threshold τ such that $R := \{\omega : x(\omega) \geq \tau, \omega \in \Omega\}$ and ρ be an arbitrary deterministic tie-breaking rule. Let $X_{(n+1)} = -\infty$. Given*

$2 \leq B < n$, *RSPM with* ρ *and a homogeneous eligible set* R *has* POA_a *of at most*

$$\mathbb{E}\left[X_{(1)} - X_{(2)}\mathbf{1}\left[X_{(2)} \geq \tau > X_{(B+1)}\right] - \tau\mathbf{1}\left[X_{(1)} \geq \tau > X_{(2)}\right] - \sum_{e=B+1}^{n} X_{(e-B+2)}\mathbf{1}\left[X_{(e)} \geq \tau > X_{(e+1)}\right]\right].$$

Its proof structure follows that of Theorem 3.5.1, but is more involved due to the randomness in the mechanism. Formal proof is deferred to Appendix B.6.16. This result depicts the trade-off between the budget and the the efficiency of a mechanism. Indeed, the upper bound on POA presented above decreases as B increases, and by plugging $B = n$ we can exactly recover the guarantee provided in Theorem 3.5.1. This is fairly intuitive as if the principal can evaluate more candidates, then she can commit to more efficient solutions. Similar to Remark 3.4.8, RSPM can be either BM or PIM depending on the choice of the eligible sets. For PIM, we can directly obtain the following result.

Corollary 3.7.6. *There exists a PIM having budget* B *with* POA_a *of* $\mathbb{E}\left[X_{(1)}\right] - \mathbb{E}\left[X_{(n-B+2)}\right]$.

Importantly, quantifying the additive approximation factor in Corollary 3.7.6 is intrinsically more challenging than the previous result without the examination cost, since we need to deal with the expected difference of two order statistics which are not consecutive. One straightforward approach would be to simply sum up the bound for two consecutive ones, but we found this bound to be cumbersome and couldn't find any take-away. Analyzing or simplifying this gap for arbitrary budget B and deriving analogous results as in Section 3.5 remains a major open problem.

Although RSPM has the benefit of its simplicity, however, it is not a clever choice as it examines the candidates in uniformly at random. A better approach might be to exploit the prior information on the distributions regarding the quality of candidates, which might be an interesting topic.

Part II

Delegation of Production

In this part of the article, I explore the strategic interaction between content providers in online contents platforms and the platform's recommendation algorithm.

Recommendation algorithms are ubiquitous in online platforms such as video streaming services, e-commerce sites, and social media. These algorithms aim to increase user engagement by recommending content that users are most likely to prefer, based on estimates of user preferences and content features. The creation of this content is typically handled by individuals known as content producers. However, producing high-quality content requires significant time and effort. To keep content producers engaged in the ecosystem, platforms often provide direct monetary compensation or encourage them to embed advertisements in their content to generate third-party revenue. Crucially, the incentives for content producers are closely tied to the recommendation algorithm itself, as greater visibility to users translates to higher potential earnings.

Despite the strategic nature of content production, most recommendation algorithms do not account for the producers' strategic behavior—namely, their efforts to maximize revenue while managing production costs. This interaction is critical to designing an efficient ecosystem, yet it remains underexplored in existing models.

Chapter 4: Optimal Contest for Online Platforms

This chapter presents the results of [20] (STOC 2026).

4.1 Introduction

The theory of *contest design*, introduced nearly half a century ago by [80], has been one of the most central problems in game theory and mechanism design, attracting attention traditionally from economists [4, 81–83] to more recent interest from computer science [84–87], with numerous applications including research contests, sports tournaments, crowdsourcing, and online content or question-answering platforms. In the contest design problem, there are n strategic contestants, each of whom decides the amount of costly effort to exert in the competition, which maps to their quality q given a cost function $c(q)$ to maximize their payoff. The contest designer aims to design a rank-based mechanism that maps each rank i (based on quality) of each contestant to an allocated budget p_i for $i \in [n]$, given a fixed total (normalized) budget $\sum_{i\in[n]} p_i = 1$.

As an illustrative example, consider the problem of online content platforms or social media that design a recommendation algorithm to decide which content to deliver to each user. The production of content on these platforms is typically handled by individuals known as content producers. However, producing high-quality content requires significant time and effort. To keep content producers engaged in the ecosystem, platforms provide monetary compensation to encourage the production of high-quality content. Crucially, the incentives for content producers are closely tied to the recommendation algorithm itself, as greater visibility to users translates to higher potential earnings. The platform not only needs to maintain high average quality of the content registered on the platform, but also guarantee user welfare, *i.e.,* the quality of content that is actually shown to the user.[1]

The seminal result by [4] reveals that the optimal mechanism uniformly distributes the bud-

[1]Looking forward, a convex combination of these two metrics turns out to be neither convex nor concave, which primarily motivates our study of such generalized objective functions. Still, our results accommodate far more than this specific application, as will be discussed shortly.

get among the contestants except the last-ranked one (henceforth, the UNIFORMBUTLAST policy) when the cost function is linear and the objective function is the average quality among the contestants. Since then, despite significant interest, there have been only a few known results even under simplistic cases [81–83, 86], stating that either of the two extreme mechanisms, UNIFORMBUTLAST or HARDMAX (which allocates all the budget to the highest-ranked contestant), is optimal in some special cases. In particular, most of these papers assume a linear objective function to maximize the average quality of the contestants [4, 82, 85, 86], or a convex (concave) function over the quality even in the most general case, with a correspondingly restrictive class of cost functions [83]. In practice, however, many scenarios, such as recommender system contests, involve objective functions that are much more complicated.

In this context, the following major open questions remain unresolved, with no partial evidence having emerged for decades.

Q1: *Would* HARDMAX *or* UNIFORMBUTLAST *still be optimal beyond convex objectives?*

Q2: *If not, what would the optimal mechanism be, and is it even structured?*

In this paper, we answer these questions both negatively and affirmatively. First, we show that neither policy is optimal in general, but that a mechanism interpolating between the two is optimal, *i.e.,* one that assigns potentially high prize to the first rank (similar to HARDMAX) but uniformly distributes the remaining prizes to intermediate ranks except the last (similar UNIFORMBUTLAST)—a mechanism that has never been studied in the literature. On the other hand, we show that either HARDMAX or UNIFORMBUTLAST is optimal in certain special cases, as evidenced by prior work, and we provide an almost complete characterization of the conditions under which such results hold,

thereby filling missing gaps in the literature.

To highlight our main technical contributions:

1. We first establish the existence of a symmetric mixed Nash equilibrium (MNE) for nontrivial policies[2] using topological arguments on discontinuous games from [88]. We further prove that the symmetric MNE is unique.

2. Towards characterizing the optimal structure, we show that either the HARDMAX policy (with $p_1 = 1$ and $p_2 = \ldots = p_n = 0$) or the UNIFORMBUTLAST policy (with $p_1 = \ldots = p_{n-1} = 1/(n-1)$ and $p_n = 0$) is optimal if the objective function is either user welfare (motivated by recommender systems) or the average quality of the contestants, exhibiting a stark *phase transition* depending on whether the cost function is convex or concave. In fact, we provide much more general conditions for such optimality, complementing partial evidence [4, 83, 86] in the literature.

3. We extend our analysis to the case where the objective function is an arbitrary convex combination of user welfare and average quality, leading to an optimization problem involving a sum of convex and concave components, which typically does not admit a closed-form solution. Nevertheless, we identify the *structure of the optimal policy* to be $p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$. Surprisingly, this result extends to significantly more general objective functions–arbitrary *posynomial*s over qualities–subsuming classic measures in mechanism design such as social welfare, order statistics, (inverse) S-shaped functions, and exponential utility, which have received little or no attention in the literature, to the extent of our knowledge.

4. Furthermore, thanks to the structural characterization, we obtain a *fully polynomial-time approx-*

[2] A policy is nontrivial if it is not $p_1 = \ldots = p_n$, which induces a trivial pure Nash equilibrium of zero effort.

imation scheme to compute the optimal policy, assuming the constant parameter in the cost function.[3] Interestingly, our techniques yield a *general reduction* from a structured high-dimensional nonconvex optimization to a single-dimensional optimization problem, which implies potential polynomial-time solvability for a wide class of high-dimensional non-convex optimization.

In particular, we note that our results subsume prior results in comprehensive manner: linear cost [4] and convex cost settings [86], average quality as an objective [81], restricted class of policies [82], and convex (concave) utility function with concave (convex) cost [83]. It may be also worth noting that our existential result uses topological arguments by [88] rather than characterizing several analytical properties of symmetric profiles tailored to each setting [81, 83], which we believe may be of general interest to the literature, formalizing unstructured observations in the literature.

Further, en route to our results on the structural characterization of $p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$, we observe a surprising connection between the shape of the gradient sequence of the objective function and number of transition points in optimum for a constrained symmetric function optimization problem, which may be of independent interest to broader optimization community.

In what follows, we give an overview of our results and techniques. Related works are discussed in Section 7.1.2, and we present formal problem setup in Section 7.2, followed by main technical results. Main technical results, from problem setup in Section 7.2 to main results in Section 4.7, will be presented primarily based on the recommender systems application, and we discuss generalized objective functions and their implications on Section 4.9. All the proofs are deferred to appendices.

[3] This is reasonable in practice, since the only scaling parameter in real-world scenario would be the number of contestants n whereas the cost function is fixed.

4.1.1 Results and Techniques

Before presenting a summary of our results and proof techniques, we briefly introduce our setting. More details can be found in Section 7.2. Looking ahead, our characterization of the optimal structure is summarized in Table 4.1. For ease of exposition, we present our model and results based on the scenario of recommender systems, while noting that the results are broadly applicable to various settings studied in the literature, such as research contests and crowdsourcing contests.

There are n content producers (*i.e.,* contestants), each of whom decides on a quality q_i for their content (*i.e.,* outcome). Producing content with quality q requires a cost of $c(q) = q^\beta$ for $\beta > 0$.[4] The recommender system (RS, *i.e.,* a contest designer) designs a rank-based recommendation policy $\mathbf{p} = (p_1, \ldots, p_n) \in \Delta$ where $\Delta = \{\mathbf{p} \in \mathbb{R}^n_{\geq 0} : p_1 \geq \ldots \geq p_n \geq 0, \sum_{i=1}^n p_i = 1\}$. Then, whenever a user joins the platform, the RS sorts the existing content in decreasing order of quality and recommends the k-th highest quality content with probability p_k,[5] where ties are broken uniformly at random.

Producer i receives a unit reward of 1 whenever their content is recommended. Hence, their expected reward under this policy when their content has rank k is p_k. The eventual payoff of producer i is the expected reward minus the cost of producing content with the corresponding quality, *i.e.,* $p_k - c(q_i)$ if q_i is ranked k-th given others' strategies.

The RS aims to maximize a convex combination of user welfare ($\mathcal{W}(\bar{\mu}, \mathbf{p})$) and platform quality ($\mathcal{Q}(\bar{\mu}, \mathbf{p})$) at a symmetric MNE $\bar{\mu}$. Here, user welfare denotes the expected quality of content shown to users, while platform quality refers to the average expected quality of all content submitted to the system.[6] Note that $\bar{\mu}$ represents the product distribution over producers' mixed strategies, where each

[4] All our results readily adapt to the linear transformation $c(q) = \gamma q^\beta$.

[5] From the general contest design problem's perspective, each p_k can be viewed as a prize with normalized total budget. Thus, when we write *probability*, this refers to *prize* in the context of general contest design literature.

[6] See Section 7.2 for formal definitions.

producer may randomize over content quality. The goal is to design a policy $\mathbf{p}$ that maximizes the objective: $\alpha\mathcal{W}(\bar{\mu}, \mathbf{p}) + (1-\alpha)\mathcal{Q}(\bar{\mu}, \mathbf{p})$.

We say that the RS is *single-minded* if $\alpha = 0$ or 1, *i.e.,* it is interested solely in maximizing either user welfare or platform quality, and *convex-minded* if $\alpha \in (0, 1)$, *i.e.,* when the objective function is truly a convex combination of the two. This formulation captures the trade-off between maximizing immediate user satisfaction and ensuring long-term ecosystem quality.

Existence of symmetric MNE (Section 4.4.1)

First, we show that a pure Nash equilibrium (PNE) does not exist under any policy where $p_i \neq p_j$ for some $i, j \in [n]$, thereby generalizing the non-existence result of [89] beyond the HARDMAX policy. This motivates the study of mixed strategy equilibria. Given the homogeneity of the setting, it is natural to focus on symmetric mixed Nash equilibria (MNE), where all producers adopt the same strategy.[7]

On the other hand, standard existence theorems do not directly apply due to discontinuities in the producers' payoff functions: a small increase in content quality can cause a discontinuous jump in utility, as the recommendation probability may shift abruptly from p_{i+1} to p_i for some i.

To overcome this challenge, we exploit the seminal result of [88], which establishes the existence of symmetric MNE in quasi-symmetric games under two key conditions: diagonally better-reply security and upper semi-continuity of the payoff function with respect to the mixed strategy. We verify that our setting satisfies these conditions by metrizing the space of mixed strategies (endowed with the weak* topology) using the Lévy-Prokhorov metric under the ℓ_2 norm.[8] This approach allows us to reason about distances in the space of pure strategies rather than directly analyzing the underlying

[7] Symmetric equilibrium is widely adopted, especially in symmetric games, due to its simplicity and anonymity: players do not need to coordinate with others about which strategies to play in order to reach equilibrium outcomes.

[8] We refer to Section 4.4 and Appendix D.1.2 for more details on the weak* topology and the Lévy-Prokhorov metric.

topological properties of the space of distributions.

Our existential result generalizes (i) that of [89], who prove the existence of symmetric MNE when the platform runs the HARDMAX policy, to a broader class of policies, and (ii) that of [81], who characterize properties of the symmetric MNE for the linear cost function. In particular, the existential proof by [81] relies heavily on the necessary conditions that an equilibrium must satisfy if it exists and manually argues that a system of equations provides a symmetric MNE with bounded support and that any deviation beyond the support only decreases the payoff, all of which are largely tailored to their specific setting. We provide a more generic proof based on topological arguments, which easily extends beyond the linear cost function case in [81].[9]

Setting	β	**Techniques**	**Optimal Policy**
$\alpha = 1$ (User welfare)	Any	Schur-convex	HARDMAX
$\alpha = 0$ (Platform quality)	< 1	Schur-convexity	HARDMAX
	$= 1$	Neutral (linear)	Any
	> 1	Schur-concavity	UNIFORMBUTLAST
$\alpha \in (0, 1)$ (Mixed)	Any	Variation diminishing	$p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$
Arbitrary posynomial	–	Variation diminishing	$p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$

Table 4.1: Optimal policy under different values of α (objective) and β (cost curvature). "Any" in the optimal policy column denotes that any policy with $p_n = 0$ would induce the same objective values.

Uniqueness of symmetric MNE (Section 4.4.2)

Recall that our existential theorem guarantees that, for any nontrivial policy $\mathbf{p}$, there exists at least one corresponding symmetric equilibrium. Looking ahead to Section 4.5, we prove that the RS's optimization problem at any symmetric MNE, given a policy $\mathbf{p}$, can be reduced to a simpler optimization problem that involves only $\mathbf{p}$ in the objective and constraints—without requiring $\boldsymbol{\mu}$. This

[9] [83] informally claim that the existential argument by [81] easily extends beyond the linear cost case without formal justification, which we believe may need to be formally justified given their subtle arguments. In any case, we believe our results can serve as a formal framework to argue the existence of symmetric MNE in the contest literature.

suggests that the symmetric MNE may be unique. Indeed, in Theorem 4.4.5, we prove that this is the case; that is, the symmetric MNE corresponding to a nontrivial policy is unique. This follows from some observations on the characteristics of the cumulative distribution function of the symmetric MNE, stemming from the properties of the equilibrium.

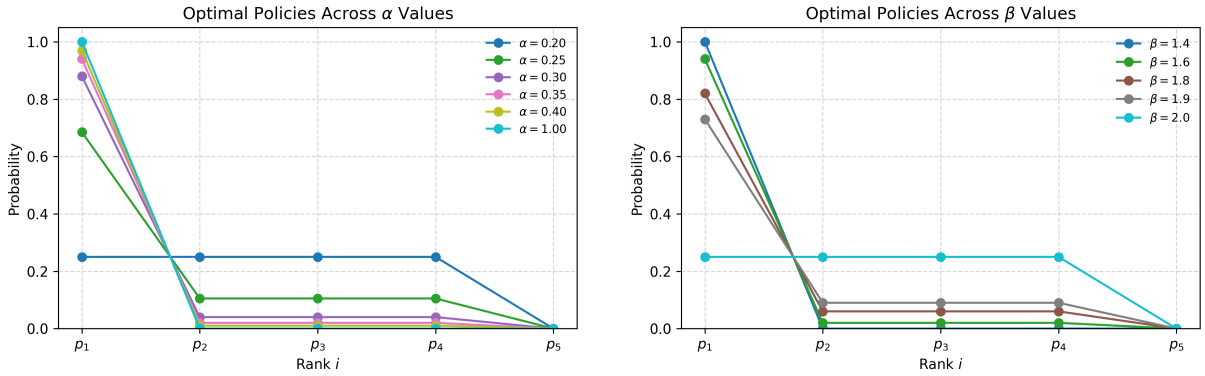
Optimal rank-based policy: optimization over Bernstein basis polynomials (Section 4.5)

We now turn to the problem of characterizing the optimal rank-based recommendation policy, assuming content producers play according to the symmetric MNE. In this setting, the RS aims to solve the following optimization problem:

$$(\text{OPT}_1) \quad \underset{\mathbf{p}\in\Delta}{\text{maximize}} \quad \alpha\mathcal{W}(\boldsymbol{\mu},\mathbf{p}) + (1-\alpha)\mathcal{Q}(\boldsymbol{\mu},\mathbf{p})$$
$$\text{subject to} \quad \boldsymbol{\mu} \text{ is a symmetric MNE}$$

At first glance, solving this problem appears intractable because it requires an explicit characterization of the equilibrium distribution $\boldsymbol{\mu}$.

Interestingly, we show that solving the RS's problem does not require explicitly computing $\boldsymbol{\mu}$. We first prove that, in any optimal policy, the last-ranked content should receive zero exposure, *i.e.*, $p_n = 0$ using correlation inequality, and further that the symmetric MNE does not have a point mass. Building on this, we derive a reduced optimization problem using characteristics of symmetric MNEs and algebraic properties of Bernstein basis polynomials along with their intricate relation to the platform's objective function, stated as follows.



(a) Optimal Policy with respect to α

(b) Optimal Policy with respect to β

Figure 4.1: Structures of optimal policy with respect to different α and β values given $n = 5$ producers. We compute the optimal policy using a brute-force search over every possible policy after discretizing the policy space with granularity 0.005, and using trapezoidal rule to approximate integral values over the grid $[0, 0.001, 0.002, \ldots, 1]$. For the left figure, we set $\beta = 2$ and for the right figure, we set $\alpha = 0.24$. This yields that our characterization is tight in a sense that the structures $p_1 > p_2 = \ldots = p_{n-1} \geq p_n = 0$ or $p_1 = p_2 = \ldots = p_{n-1} \geq p_n = 0$ indeed appear in the optimal policies.

Theorem 4.1.1. *The platform's problem can be written as the following optimization problem:*

$$\begin{aligned} (\textsc{Opt}) \quad \underset{\boldsymbol{p}}{\textit{maximize}} \quad & \alpha n \int_0^1 h(x,\boldsymbol{p})^{1+1/\beta} dx + (1-\alpha) \int_0^1 h(x,\boldsymbol{p})^{1/\beta} dx \\ \textit{subject to} \quad & p_1 \geq p_2 \geq \ldots \geq p_n = 0 \\ & \sum_{i=1}^{n} p_i = 1 \end{aligned}$$

Note that $h(x, \mathbf{p}) = \sum_{i=1}^{n} a_i(x) p_i$, and $a_i(x) = \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}$ are the $(n-1)$-dimensional Bernstein basis polynomials.[10] This reduction reveals a key insight: *we can optimize over* $\boldsymbol{p}$ *without explicitly solving for the equilibrium* $\bar{\mu}$, enabling a tractable analysis of the RS's design problem. Further, this optimization result immediately implies an interesting corollary (Proposition 4.5.1), stating that HARDMAX can be arbitrarily worse than the optimal policy as the number of producers n increases.

[10]The standard i-th n-dimensional Bernstein basis polynomial is $b_{i,n}(x) = \binom{n}{i} x^i (1-x)^{n-i}$.

Single-minded case via Schur-convexity (Section 4.6)

Despite the simplified formulation, solving the optimization problem remains challenging, as it involves a sum of two integrals—each defined over a nontrivial function $h(x, \mathbf{p})$ raised to a power. A key analytical feature is that both integrands share the form $h(x, \mathbf{p})^r$, where $h(x, \mathbf{p})$ is the inner product between the Bernstein basis polynomials and the policy vector $\mathbf{p}$.

To make progress, we first analyze the properties of $h(x, \mathbf{p})^r$ for $r \geq 0$. Interestingly, if we extend the domain of $h(x, \mathbf{p})$ to the full n-dimensional probability simplex—by sorting any input vector in decreasing order before applying h—the resulting symmetric function becomes *Schur-convex* when $r \geq 1$ or $r \leq 0$, and *Schur-concave* otherwise.[11]

This characterization immediately yields several results. When $\alpha = 1$, the RS's objective becomes Schur-convex, as $r = 1 + 1/\beta > 1$, which—together with the constraint $p_n = 0$—implies the optimality of the HARDMAX policy. When $\alpha = 0$, the objective is Schur-convex if $\beta \leq 1$ and Schur-concave if $\beta \geq 1$, implying that HARDMAX is optimal under concave costs, and UNIFORMBUTLAST under convex costs, *i.e.,* a phase transition occurs. Notably, when $\beta < 1$, both user welfare and platform quality are Schur-convex, reinforcing the optimality of HARDMAX. This is formalized in the following theorem.

Theorem 4.1.2. *Depending on the weight parameter α and the cost parameter β, the following holds:*

1. *If $\alpha = 1$ or $\beta \leq 1$, or $\alpha = 0$ and $\beta < 1$, then* HARDMAX *is optimal.*

2. *If $\alpha = 0$ and $\beta > 1$, then* UNIFORMBUTLAST *is optimal.*

3. *If $\alpha = 0$ and $\beta = 1$, then any $\boldsymbol{p} \in \Delta$ with $p_n = 0$ is optimal.*

[11] A function $f : \mathbb{R}^n \to \mathbb{R}$ is Schur-convex (resp. Schur-concave) if $f(x) \geq f(y)$ (resp. $f(x) \leq f(y)$) whenever x majorizes y. See Section 4.6 for details.

A similar result is often observed in the literature [4,83], *i.e.,* either of these two extreme policies is optimal in specific scenarios. Our analysis via Schur-convexity provides clearer intuition on why these policies turn out to be optimal in these regimes, with broader generalizability due to its analytical clarity. For instance, this approach allows the extension to several generalized objective functions discussed in Section 4.9, combined with our techniques for the convex-minded case.

Convex-minded case via variation diminishing property (Section 4.7)

In the general case where $\alpha \in (0, 1)$, the objective becomes a combination of a Schur-convex and a Schur-concave function. This poses a significant analytical challenge, as such objective functions typically do not admit a closed-form characterization of the optimal solution; *e.g.,* consider optimizing a function that is the sum of a convex and a concave function.

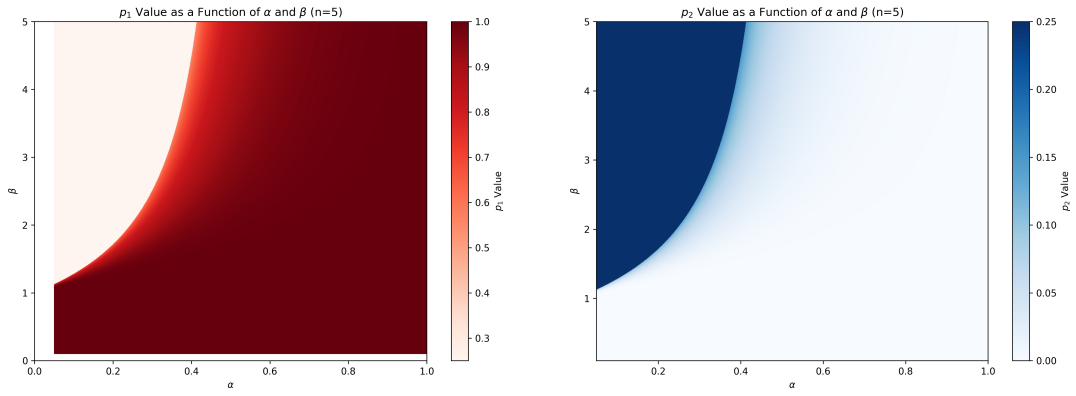
Despite this difficulty, we establish that the optimal policy admits a remarkably structured form: $p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$. In other words, the optimal $\mathbf{p}$ is a piecewise constant vector with at most two distinct values (possibly differing at the top and bottom ranks), which is formally stated as follows:

Theorem 4.1.3. *For any $\alpha \in [0, 1]$ and $\beta > 0$, any optimal policy has the structure:*

$$p_1 \geq p_2 = p_3 = \ldots = p_{n-1} \geq p_n = 0.$$

To derive this result, we develop several technical tools involving Bernstein basis polynomials and their matrix representations, which may be of independent interest.

Formally, we show that for any real numbers $0 < x_1 < \ldots < x_k < 1$ and integers $0 \leq i_1 < \ldots < i_k \leq n$, the matrix $(a_{i_s}(1 - x_l))_{s,l\in[k]}$ exhibits a desirable property known as *total positivity*,



(a) p_1 of optimal policy with respect to α and β (b) p_2 of optimal policy with respect to α and β

Figure 4.2: Structures of optimal policy with respect to different α and β values given $n = 5$ producers. Darker red means larger p_1 values, and darker blue means larger $p_2 = p_3 = \ldots = p_{n-1}$ values. Each of the optimal policy is computed in a brute force manner after discretizing policy space with granularity 0.001 and using trapezoidal rule with number of points 200. Parameters α and β are selected from the ranges $[0.05, 1]$ and $[0.1, 5]$, respectively, with 1000 points calculated for each.

as introduced by [90]—meaning that every minor of the matrix has a positive determinant. This is established by expressing each determinant as a generalized *Vandermonde* matrix [91].[12]

We then leverage a seminal result from the theory of total positivity and its connection to the *variation diminishing property*[13] [90, 92] to show that the partial derivatives of the objective function with respect to p_i for $i \in [n-1]$ exhibit a *quasiconvex* structure.[14] That is, these partial derivatives decrease up to a certain index and then increase, as a consequence of bounding the number of sign changes in these sequences by those of a related quasiconvex function involving $h(x, \mathbf{p})$. This is the key ingredient of our main results, formally stated in Lemma 4.7.2. In particular, it provides a sufficient condition on the objective function under which such properties hold, which will easily allow us to generalize the results, as will be discussed shortly.

[12] See Section 4.7 for technical details.

[13] The variation diminishing property refers to the phenomenon where the number of sign changes in a function is reduced when transformed through a totally positive operator. See Section 4.7 for further explanation.

[14] We use "quasiconvex" to mean that the function first decreases and then increases with respect to the index i. See Definition 4.7.1 for the formal definition.

Finally, we formulate the Lagrangian of the optimization problem and apply the *Karush–Kuhn–Tucker (KKT)* conditions to derive the necessary optimality conditions, leveraging the geometric insights obtained from the structure of the partial derivatives.

As a final remark, Figure 4.2 illustrates how the values of p_1 (left) and p_2 (right) vary with α and β. Combined with the fact that HARDMAX can be arbitrarily worse than the optimal policy as n increases, these results suggest that allocating nonzero probabilities to lower-ranked content is crucial for designing optimal recommendation policies in the presence of strategic content producers.

Implication on high-dimensional nonconvex optimization problem. In fact, our techniques for such structural characterization yields an interesting implication for a general optimization problem. Specifically, given a monotone increasing nonnegative (possibly nonconvex) function f and a constant $C > 0$, consider the following generic optimization problem:

$$\begin{aligned} (\text{GEN}) \quad & \underset{\mathbf{x}\in\mathbb{R}^n}{\text{maximize}} \quad f(\mathbf{x}) \\ & \text{subject to} \quad x_1 \geq x_2 \geq \ldots \geq x_n \geq 0 \\ & \qquad\qquad \sum_{i=1}^{n} x_i = C \end{aligned}$$

Then, a general version of our structural result can be written as the following:

Corollary 4.1.4. *Consider the constrained optimization problem (*GEN*). Suppose that the gradient of f with respect to each x_i can be written as the following:*

$$\frac{\partial f(\boldsymbol{p})}{\partial x_i} = \int_D w(y, \boldsymbol{p}) g_i(y) dy,$$

for a continuous bounded interval D, a function $g_i : D \to \mathbb{R}_{\geq 0}$ for $i \in [n]$, and another function $w(y, \boldsymbol{p}) : D \times \mathbb{R}^n \to \mathbb{R}$. Assume further that $g_i(y)$ is STP$_n$,[15] i.e., *strongly totally positive of order n, and $w(y, \boldsymbol{p})$ is quasiconvex over $y \in D$,* i.e., *it decreases then increases, given any $\boldsymbol{p}$. Then, it follows that the optimum of f satisfies $x_1 \geq x_2 = \ldots = x_n$.*[16]

This structural characterization has a significant implication. Since once we know that $x_1 \geq x_2 = \ldots = x_n = x$, due to the monotonicity of f, it is immediate that $x_1 = C-(n-1)x$. Thus, the only variable that the algorithm needs to decide is $x \in [0, C/(n-1)]$, *i.e.,* it becomes a single-dimensional optimization problem on a bounded interval. Since single-dimensional optimization usually admits an efficient tractable solution under reasonable assumptions, this implies that the optimum for the original high-dimensional nonconvex optimization can be efficiently solved. Indeed, in what follows, we discuss that the optimal policy can be efficiently computed for our objective function in (OPT) in Theorem 4.1.1.

Computing the optimal policy (Section 4.8)

Note that, given the structural results, the optimization problem reduces to a single-variable optimization. However, the objective function is still a combination of convex and concave components, which may not always admit a tractable solution. In Section 4.8, we present an efficient algorithm that computes an approximately optimal policy with either additive or multiplicative error ε, running in $\mathrm{poly}(n, 1/\varepsilon)$ time given a constant β. The algorithm assumes access to a value oracle that returns the objective function value for any input.[17] It works by iteratively narrowing an interval using a care-

[15] We refer to Definition 4.7.4 for more details.

[16] Indeed, for our objective function presented in (OPT), the proof of Lemma 4.7.2 asserts that this subsumes our result by setting $w(x, \mathbf{p}) = n\alpha(1 + 1/\beta)h(x, \mathbf{p})^{1/\beta} + (1 - \alpha)/\beta h(x, \mathbf{p})^{1/\beta - 1}$ and $g_i(x) = a_i(x)$.

[17] We further discuss that the access to the value oracle can be removed by approximating the integrals in the objective function while keeping the order of error rate.

fully constructed pair of upper and lower bounds on the objective, whose gap shrinks as the interval length decreases. This is possible because the integrand in the reduced objective function is an affine transformation of the single variable p_1, followed by exponentiation parameterized by β. Using this structure, a branch-and-bound-style algorithm is employed to explore the search space: it branches on subintervals when necessary and prunes those that violate optimality conditions, efficiently converging to a near-optimal solution.

Beyond RS: general objective functions (Section 4.9)

Finally, we discuss how our proposed techniques accommodate a broader range of objective functions in Section 4.9. First, we show that an arbitrary posynomial[18] function over the produced contents' qualities at the symmetric MNE still admits the exact same optimal structure:

Theorem 4.1.5. *Consider the following generalized objective function at symmetric MNE $\bar{\mu}$ given $\boldsymbol{p}$:*

$$\mathcal{G}(\bar{\mu}, \boldsymbol{p}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{q_i \sim \mu_i} \left[e_m q_i^{k_m} + e_{m-1} q_i^{k_{m-1}} + \ldots + e_1 q_i^{k_1} \right],$$

where $q_i \sim \mu_i$ for a symmetric strategy profile μ_i in $\bar{\mu}$. For any $\beta > 0$, suppose e_j's and k_j's satisfy the following property:

1. *There exists $j^* \in [m]$ such that $e_j(k_j - \beta) \leq 0$ for $j \leq j^*$ and $e_j(k_j - \beta) \geq 0$ for $j > j^*$.*
2. *All the elements in the sequence $e_j(k_j - \beta)$ for $j \in [n-1]$ is nonnegative or nonpositive.*

Then, any optimal policy follows

$$p_1 \geq p_2 = p_3 = \ldots = p_{n-1} \geq p_n = 0.$$

[18] A posynomial function is a polynomial with nonnegative coefficients whose exponents may be noninteger.

As a special case, if e_j's are all nonnegative, this condition follows.

The proof is similar to Theorem 4.1.3, *i.e.,* we show that the objective function satisfies the condition in Lemma 4.7.2. On the other hand, this condition is established using different approaches, by proving how the coefficients of an ordered posynomial relate to the quasiconvexity of the objective function via a generalized version of Descartes' rule of signs [93, 94].

Notably, this generalized argument implies several interesting corollaries. For instance, our results apply to the following objective functions, thereby admitting the same noted optimal structure:

1. Social welfare, which sums up every involved player's utility function combined with a generalized platform utility (posynomial).

2. Inverse S-shaped functions, *e.g.,* $2q^3-3q^2+2q$, which are concave-then-convex over the induced contents' qualities, capturing a smoothed version of thresholded utility.[19]

3. Largest order statistics, corresponding to the highest quality among the contents produced by the producers.

4. Exponential utility $e^{\lambda q}$ for any $\lambda > 0$ and even arbitrary convex combination among them.

Remarkably, our results allow many positive combinations of these objective functions, as long as the sequence of the objective function's partial derivatives with respect to the p_i's remains quasiconvex, which will become clear from our main analysis.

[19] These are the opposite of S-shaped functions from prospect theory [95] in behavioral economics, which exhibit convex-then-concave geometry.

4.2 Related Works

Our model is closely related to the literature on rank-based contests, where the allocation probability p_i for the i-th highest-quality content parallels the assignment of a prize p_i to the i-th highest-ranked contestant in a competition with n participants.[20] For a comprehensive survey of classic work on contest design—including rank-based and Tullock contests—we refer to [96].

Rank-based contests. The design of rank-based contests—where prizes depend on contestants' ranks—has been extensively studied since [80] and [4]. [4] analyze optimal prize allocation under linear costs and common risk preferences, proving that the UNIFORMBUTLAST policy is optimal in a complete information setting. This aligns with our model when $\beta = 1$ and $\alpha = 0$. [82] extend this to general cost functions under incomplete information, focusing on top-two prize structures and identifying when HARDMAX (winner-take-all) is optimal. [81] examine equilibrium properties under exogenous policies with linear costs, while [86] show that uniformly rewarding the top quartile yields a constant-factor approximation for convex costs when maximizing average quality. [85] study a strategic participation model with fixed costs and qualities, showing that distributing equal prizes to top contestants is optimal with identical costs and that HARDMAX offers a 3-approximation more generally. Finally, the recent work by [83] observes a general connection between the competitiveness of the contest and monotonicity of utility function when the utility function is convex (concave) and the cost function is concave (convex), respectively, under the complete information setting.

Our model generalizes prior work by incorporating broader platform objectives and general cost functions, offering new insights into contest design under complete information. More importantly, it

[20] In the contest design literature, the total prize may not sum to 1; however, normalization allows for analogous interpretations.

is the first to study recommender systems as an application of the classic contest design in the literature.

Strategic content production in recommender system. [89] investigate competition among multiple content producers, each deciding an effort level to create content of a certain quality. Their single-dimensional model yields the same expected payoff structure as ours, though they analyze equilibria only under the HARDMAX policy. [97] examine strategic behavior by content producers who attempt to game the platform to artificially boost user engagement through clickbait. They show that content quality and gaming behavior are positively correlated, and that engagement-based optimization can reduce user welfare by incentivizing such behavior. [98] study a duopoly market where two platforms running multi-armed bandit algorithms compete for user participation. [99] explore strategic content production, where producers decide whether to participate on the platform and choose their effort levels. [100] analyze the inefficiency of top-k-style recommendation policies when content creators compete for exposure, showing that these policies can induce suboptimal equilibria. [101] challenge the conventional belief that monotone reward structures are always desirable, demonstrating that non-monotone policies can better align creator incentives and improve user welfare.

Our results complement this line of work by adopting a general and abstract model to study strategic interaction, bridging classic contest design literature with the design of recommendation algorithms.

4.3 Model

There are $n \geq 2$ content producers (henceforth, *producers*), each of whom creates a single piece of content through costly effort. Each producer i strategically chooses an effort level to generate a competitive content item of quality q_i, whose interpretation will be made clear shortly.

The produced content is registered on an online platform with a recommender system (RS). Whenever a user visits the platform, the RS recommends one of the registered content items according to a predetermined policy.

We consider a class of *rank-based recommendation policies*, where the RS commits to an allocation vector $\mathbf{p} \in \Delta$, with $\Delta = \left\{(p_i)_{i\in[n]} : p_1 \geq p_2 \geq \cdots \geq p_n \geq 0,\ \sum_{i=1}^{n} p_i = 1\right\}$, i.e., the ordered n-dimensional probability simplex. Under policy $\mathbf{p}$, the RS ranks the contents in decreasing order of quality and displays the item with the i-th highest quality with probability p_i, breaking ties uniformly at random. Formally, given the vector of content qualities $\mathbf{q} = (q_1, \ldots, q_n)$, the RS first sorts them into order statistics $q_{(1)} \geq q_{(2)} \geq \cdots \geq q_{(n)}$,[21] and then displays $q_{(i)}$ with probability p_i.

We introduce two specific policies that will appear frequently throughout the paper:

1. **HARDMAX policy**: $p_1 = 1$.

2. **UNIFORMBUTLAST policy**: $p_1 = p_2 = \ldots = p_{n-1} = \frac{1}{n-1}$ and $p_n = 0$.

In words, HARDMAX deterministically recommends the highest-quality (i.e., most preferred) content to the user. In contrast, UNIFORMBUTLAST uniformly randomly selects any content except the one with the lowest quality. These two policies represent extreme points in the policy space.

Induced game and equilibrium. Given a policy $\mathbf{p}$, we consider a game *induced by* $\mathbf{p}$ given n producers. Each producer i strategically decides the quality q_i of the content it creates from the action space $\mathbb{R}_{\geq 0}$. Each producer i may play a mixed strategy that randomizes over the action space. In this case q_i is a random variable and we write μ_i to denote the associated probability measure over $\mathbb{R}_{\geq 0}$ and $F_i : \mathbb{R}_{\geq 0} \to [0, 1]$ to denote the corresponding c.d.f. We say that a producer plays a pure strategy if μ_i has a singleton support. We write $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and $\boldsymbol{\mu}_{-i} = (\mu_1, \ldots, \mu_{i-1}, \mu_{i+1}, \ldots, \mu_n)$ to

[21] Given values $x_1, \ldots, x_n$, we denote by $x_{(i)}$ the i-th largest element (order statistic).

denote the strategy profiles except i's strategy.

Producer i's expected utility $\mathcal{U}_i$ is a function over $\boldsymbol{\mu}$ and the policy $\mathbf{p}$, which will be defined shortly. Given a policy $\mathbf{p}$, a strategy profiles $\boldsymbol{\mu}$ is called a (Nash) equilibrium if no producer has incentive to deviate from $\boldsymbol{\mu}$, *i.e.,* for any $i \in [m]$:

$$\mathcal{U}_i(\mu_i, \boldsymbol{\mu}_{-i}, \mathbf{p}) \geq \mathcal{U}_i(\mu_i', \boldsymbol{\mu}_{-i}, \mathbf{p}),$$

for any mixed strategy μ_i'. If every strategy μ_i in $\boldsymbol{\mu}$ is the pure strategy, the resulting equilibrium is pure Nash equilibrium (PNE), otherwise it is mixed Nash equilibrium (MNE). An equilibrium is called *symmetric* if every player is playing the same strategy $\mu = \mu_i$ for $i \in [n]$ under the equilibrium. In this case, we write F to denote the corresponding c.d.f. In particular, throughout, we focus on symmetric MNE of the induced game, which will formally be shown to exist in Section 4.4.[22]

Producer's utility. Producer i's expected utility consists of two components: (i) the *revenue*[23] earned when their content is recommended, and (ii) the *cost* of producing content of a given quality. Each time producer i's content is recommended, they receive a unit revenue of 1. The expected revenue $\mathcal{R}_i$ when producer i follows strategy μ_i is:

$$\mathcal{R}_i(\mu_i, \boldsymbol{\mu}_{-i}, \mathbf{p}) = \mathbb{E}_{\mathbf{q}\sim\boldsymbol{\mu}}\left[\mathbf{1}\left\{ I = i \right\}\right],$$

where I is the index of the producer whose content is recommended, based on the realized qualities $\mathbf{q}$ and the policy $\mathbf{p}$, and $\mathbf{1}\left\{ \cdot \right\}$ is an indicator variable.

[22] In fact, a pure Nash equilibrium does not always exist—even in the simple case where the platform uses the HARDMAX policy—as shown in Proposition 4.4.1.

[23] Revenue may originate from advertisements embedded by the producer or inserted by the platform (e.g., pre-roll or mid-roll ads), as well as platform-based compensation mechanisms tied to outcomes such as views, clicks, or user retention.

To generate content, a producer incurs a cost determined by the quality level. Specifically, producing content of quality q incurs cost $c(q) = q^\beta$ for some $\beta > 0$, capturing convex or concave effort costs. Thus, given others' strategies $\boldsymbol{\mu}_{-i}$, the expected utility of producer i when playing mixed strategy μ_i is:

$$\mathcal{U}_i(\mu_i, \boldsymbol{\mu}_{-i}, \mathbf{p}) = \mathcal{R}_i(\mu_i, \boldsymbol{\mu}_{-i}, \mathbf{p}) - \mathbb{E}_{q_i \sim \mu_i}[c(q_i)].$$

It is also worth noting that our model implicitly captures individual rationality (IR) condition of the producers, as playing $q = 0$ gives the deterministic utility of $p_n \geq 0$ (larger utility upon ties), and thus outperforming the outside option of not participating the contest.

User welfare. The platform's long-term success critically depends on *user welfare*, which we proxy through user engagement. We assume that when a user is recommended and consumes content of quality q, they derive utility equal to q. Given a recommendation policy $\mathbf{p}$ and the producers' strategy profile $\boldsymbol{\mu}$, we define user welfare $\mathcal{W}(\boldsymbol{\mu}, \mathbf{p})$ as the expected quality of the recommended content:

$$\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) = \mathbb{E}_{\mathbf{q} \sim \bar{\mu}} \left[\sum_{i=1}^{n} \mathbf{1}\{ I = i \} \cdot q_i \right].$$

In other words, $\mathcal{W}(\boldsymbol{\mu}, \mathbf{p})$ captures the average utility delivered to users based on the realized content qualities and the recommendation outcome. When $\mathbf{p}$ is fixed or clear from context, we write $\mathcal{W}(\boldsymbol{\mu})$ for brevity.

Platform quality. While the RS primarily aims to maximize user welfare $\mathcal{W}$ in response to strategic producer behavior, it also benefits from sustaining a high overall quality of content on the platform—a factor that can influence long-term user trust and platform reputation. We define *platform quality* as

the average quality of all registered content at the symmetric equilibrium $\boldsymbol{\mu}$ under policy $\mathbf{p}$:

$$\mathcal{Q}(\boldsymbol{\mu}, \mathbf{p}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{q_i \sim \mu_i}[q_i].$$

When $\boldsymbol{\mu}$ is symmetric, this simplifies to the expected quality under any producer's strategy. As before, we write $\mathcal{Q}(\boldsymbol{\mu})$ when $\mathbf{p}$ is clear from context.

Objective function. Finally, the RS's objective is to maximize the weighted summation of induced user welfare $\mathcal{W}(\boldsymbol{\mu})$ and the average quality $\mathcal{Q}(\boldsymbol{\mu})$:

$$\text{OBJ}(\boldsymbol{\mu}, \mathbf{p}) = \alpha \mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) + (1 - \alpha)\mathcal{Q}(\boldsymbol{\mu}, \mathbf{p}),$$

where $\alpha \geq 0$ governs the importance between the user welfare and the platform's quality.

Essentially, we are interested in a policy $\mathbf{p}$ that maximizes the objective function when the producers are playing symmetric equilibrium strategies $\boldsymbol{\mu}$.

Thus, we aim to find a policy $\mathbf{p}$ that solves the following optimization problem:

$$\begin{aligned} (\textsc{Opt}) \quad & \underset{\mathbf{p} \in \Delta}{\text{maximize}} \quad \alpha \mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) + (1 - \alpha)\mathcal{Q}(\boldsymbol{\mu}, \mathbf{p}) \\ & \text{subject to} \quad \boldsymbol{\mu} \text{ is a symmetric equilibrium} \end{aligned}$$

We say the RS is *single-minded* if $\alpha = 0$ or 1, *i.e.,* it aims to solely maximize either the user welfare or the platform quality, and *convex-minded* if $\alpha \in (0, 1)$. One might notice that in the optimal policy, it is required to have $p_i \neq p_j$ for some $i, j \in [m]$ since otherwise the dominant strategy is to play $q = 0$. Thus, we focus on the class of *nontrivial* policy in Δ such that $\mathbf{p} \neq (1/n, 1/n, \ldots, 1/n)$.

4.4 Existence and Uniqueness of Symmetric MNE

In this section, we provide our results on the existence and uniqueness of symmetric mixed Nash equilibrium. Since the platform's objective function depends on the equilibrium point, it is crucial to analyze the equilibrium point. Before presenting our main results, we first prove that there exists no PNE for any nontrivial policy, the proof of which is in Appendix D.2.2.

Proposition 4.4.1. *There exists no PNE for any nontrivial policy* $\boldsymbol{p}$.[24]

4.4.1 Existence

Proposition 4.4.1 motivates us to study the MNE instead of PNE. Specifically, we will focus on a symmetric MNE. This is because it admits tractable analysis as well as allows producers to simply commit to the strategy without considering others' potentially unorganized behavior. For instance, as the game faced by the producers are symmetric, playing a symmetric strategy is their best bet unless they coordinate to play asymmetric strategies.

Theorem 4.4.2. *For any nontrivial policy* $\boldsymbol{p}$*, the induced game (defined in Section 7.2) admits a symmetric mixed Nash equilibrium.*

The proof is not immediate due to the discontinuous nature on payoff function. We provide the proof sketch of Theorem 4.4.2 at the end of this section.

Given the existence of the symmetric MNE, the following proposition states that we can effectively wipe out the tie-breaking scenario.

[24]Note that this result strictly generalizes the nonexistence of PNE by [89] (Proposition 1) under HARDMAX rule.

Proposition 4.4.3. *Given a nontrivial policy $\boldsymbol{p}$ and symmetric MNE $\boldsymbol{\mu}$, then the corresponding c.d.f. F of $\bar{\mu}$ is atomless and strictly increasing in some interval $[0, q_{\max}]$.*

We refer to Appendix D.2.3 for the proof. This implies that the c.d.f. F of the symmetric MNE is differentiable almost everywhere (a.e.) by Lesbesgue's theorem, once it exists.

Finally, using the derived properties of the symmetric MNE, we can obtain explicit formulas for the user welfare, platform quality, and producer's payoff as follows.

Lemma 4.4.4. *Consider a symmetric MNE $\bar{\mu} = (\mu, \dots, \mu)$ given a policy $\boldsymbol{p}$. Producer i's payoff of playing q in the support of μ given others playing $\bar{\mu}_{-i}$ can be written as:*

$$\mathcal{U}_i(q, \bar{\mu}_{-i}, \boldsymbol{p}) = \left(\sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i} (1 - F(q))^{i-1} \right) - c(q),$$

where $F(\cdot)$ is the c.d.f. of μ. Further, the user welfare and the platform's quality can be written as:

$$\mathcal{W}(\bar{\mu}, \boldsymbol{p}) = \sum_{i=1}^{n} p_i \mathbb{E}_{\boldsymbol{q} \sim \bar{\mu}} \left[q_{(i)} \right], \qquad \mathcal{Q}(\bar{\mu}, \boldsymbol{p}) = \mathbb{E}_{q \sim \mu} \left[q \right],$$

where $q_{(i)}$ denotes the i-th largest order statistics among q_i's.

4.4.1.1 Proof Sketch of Theorem 4.4.2

Intuitively, the symmetry of the game—namely, the symmetry in both the providers' payoffs and their action spaces—supports the existence of a symmetric equilibrium. However, the formal proof requires a subtle argument based on the notion of *better-reply security* and topological reasoning. This is due to the fact that, while the action space is continuous, the payoff function is discontinuous over pure strategies, owing to the thresholded nature of the rank-based policy. Specifically, increasing

content quality can lead to a discontinuous jump in revenue—from p_n to p_{n-1}, and so on, up to p_1.

To prove the existence of a symmetric mixed-strategy Nash equilibrium (MNE) for any nontrivial policy $\mathbf{p}$, we rely on the framework developed by [88], which extends equilibrium existence results to discontinuous games using the notion of *better-reply security*. Our argument generalizes [89], who focus solely on the HARDMAX policy, to a broader class of rank-based recommendation policies.

Strategy Space and Topology. Each content producer chooses a quality level $q_i \in \mathbb{R}_{\geq 0}$ and may randomize over this action space. Thus, the strategy space of producer i is the space of probability measures over $\mathbb{R}_{\geq 0}$. We restrict attention to probability measures with compact support in $[0, B]$ for some sufficiently large B, which can be justified by utility bounds and standard truncation arguments as can be seen in the proof.

We endow this space of probability measures with the *weak* topology*—that is, the coarsest topology on the space of probability measures on X such that the maps $\mu \mapsto \int_X f d\mu$ are continuous for every bounded continuous function $f : X \mapsto R$. Formally, a sequence of probability measures $\{\mu_n\}$ converges weakly to μ if and only if for every bounded continuous function f, we have:

$$\int f(x)\, d\mu_n(x) \to \int f(x)\, d\mu(x).$$

This notion of convergence is crucial because it ensures a player's expected payoff is a continuous function of their opponent's mixed strategy. Specifically, if an opponent's strategy sequence $\{\mu_n\}$ converges weakly to μ, our expected payoff also converges.

Given that the producer's action space is compact, this continuity of the payoff mapping allows us to establish the existence of a symmetric MNE once we prove that the game satisfies better-reply secureness by [88].

Discontinuity and Better-Reply Security. The utility function of each producer i is discontinuous in the joint strategy profile due to tie-breaking in the ranking mechanism. That is, small changes in others' strategies can cause jumps in whether i's content is ranked j-th or $(j+1)$-th, which in turn affects the allocation probability p_j.

Despite this discontinuity, we prove that the induced game satisfies *diagonal better-reply security*, a sufficient condition for the existence of a mixed-strategy equilibrium established in [88]. The key idea is to show that for every symmetric strategy profile $(\mu, \ldots, \mu)$ with payoff u that is not an equilibrium, there exists an open neighborhood B of μ such that a producer can secure a payoff strictly greater than u along the diagonal on B, *i.e.,* when all other producers play symmetric strategies within B.

These properties allow us to invoke Corollary 5.3 of [88] (or see Theorem 4.1) to establish the existence of a symmetric mixed-strategy equilibrium.

Formal definitions and proofs are provided in Appendix D.2.1.

Remark. For completeness, we provide background on the weak* topology, probability measures on Hausdorff spaces, and technical lemmas used in the proof in Appendix D.1.2.

4.4.2 Uniqueness

In fact, we show that the symmetric MNE is even unique.

Theorem 4.4.5 (Uniqueness of the symmetric MNE)**.** *Fix $n \geq 2$, $\beta > 0$, and a nontrivial rank-based policy $p = (p_1, \ldots, p_n)$ with $p_1 \geq \cdots \geq p_n \geq 0$ (not all equal). Then the induced game admits a* unique *symmetric mixed Nash equilibrium.*

The proof uses the atomlessness and strict monotonicity of the c.d.f. of the symmetric MNE

shown in Proposition 4.4.3, along with several observations that each producer's payoff function except the cost term is strictly increasing and continuous on $F(q)$, thereby admitting a continuous inverse function. Then, the symmetric MNE becomes unique if its support is the same, which can be rather easily verified from the payoffs at the boundary points.

Remark 4.4.6. Notably, thanks to our characterization on properties of the symmetric MNE, we further observe that the producers can easily play such equilibrium strategies. Formally, they can sample $U \sim \text{Uniform}[0,1]$, and play $q = h(U, \mathbf{p})^{1/\beta}$ where the function h is defined in (4.1). This follows from the relation (D.18) combined with the correspondence between sampling quantiles uniformly at random and sampling q directly from F.

4.5 Detaching Equilibrium Dependency in Optimization

Recall that the RS's objective, denoted as OPT, involves optimizing over the policy $\mathbf{p}$ while accounting for the induced symmetric MNE. In this section, we show that this problem can be reformulated as a pure optimization over $\mathbf{p}$, without explicitly involving the equilibrium strategies.

Our key observation is that, under a symmetric MNE, any pure strategy in the support must yield the same expected utility, given that all other producers follow the equilibrium strategy. Leveraging this property, we first establish that any optimal policy must satisfy $p_n = 0$ using correlation inequality. Then, through algebraic manipulation, we express both user welfare $\mathcal{W}(\bar{\mu}, \mathbf{p})$ and platform quality $\mathcal{Q}(\bar{\mu}, \mathbf{p})$ entirely in terms of $\mathbf{p}$, eliminating explicit dependence on $\bar{\mu}$. Interestingly, the resulting expressions for $\mathcal{W}$ and $\mathcal{Q}$ share the same functional form, differing only in the exponent—highlighting a unified structure in the RS's objective.

To present our main results, we introduce two notations for ease of exposition.

$$a_i(x) = \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}, \qquad h(x, \mathbf{p}) = \sum_{i=1}^{n} a_i(x) p_i. \tag{4.1}$$

Note that $a_i(x)$ is exactly the *Bernstein basis polynomial* of degree $n-1$. The standard Bernstein polynomial $b_{\nu,n}$ is defined as: $b_{\nu,n}(x) = \binom{n}{\nu} x^{\nu}(1-x)^{n-\nu}$, for $\nu = 0, \ldots, n$. Due to our construction of $a_i(x)$, observe that for each $i \in [n]$, we have: $a_i(x) = b_{n-i,n-1}(x)$.

Now, our main theorem on the reduced optimization formulation can be stated as follows:

Restatement of Theorem 4.1.1. *The platform's problem can be written as the following optimization problem:*

$$\begin{aligned} (\textsc{Opt}) \quad \underset{\boldsymbol{p}}{\textit{maximize}} \quad & \alpha n \int_0^1 h(x, \boldsymbol{p})^{1+1/\beta} dx + (1-\alpha) \int_0^1 h(x, \boldsymbol{p})^{1/\beta} dx \\ \textit{subject to} \quad & p_1 \geq p_2 \geq \ldots \geq p_n = 0 \\ & \sum_{i=1}^{n} p_i = 1 \end{aligned}$$

We write $G(\boldsymbol{p})$ to denote the objective function in the above optimization.

Essentially, we derive algebraic connections between the objective function and the Bernstein basis polynomial as well as its derivatives. Then, we write a condition for a symmetric strategy profiles to become an equilibrium. Manipulating the derived equation and carefully combining with the platform's original objective function, we derive the reduced optimization formulation without introducing the equilibrium itself. The proof is deferred to Appendix D.3.1 along with other necessary technical lemmas.

4.5.1 HARDMAX can be Arbitrarily Bad

An immediate implication of Theorem 4.1.1 is that in the single-minded case of $\alpha = 0$ with convex cost function $\beta > 1$, HARDMAX can be arbitrarily worse than the optimal policy as n increases.

Proposition 4.5.1 (Price of HARDMAX policy)**.** *Consider convex cost function of $\beta > 1$ and the RS maximizes average quality,* i.e., $\alpha = 0$. *Let $G(\boldsymbol{p})$ the objective function obtained in Theorem 4.1.1,* i.e., *if $\alpha = 0$ then $G(\boldsymbol{p}) = \int_0^1 h(x, \boldsymbol{p})^{1/\beta}$. Then, for any $\varepsilon > 0$, there exists a sufficiently larger n such that* HARDMAX *policy becomes arbitrarily bad than the optimal policy,* i.e., $G(\text{HARDMAX})/G(\text{OPT}) < \varepsilon$.

In particular, the proof follows from comparing HARDMAX with UNIFORMBUTLAST, which is deferred to Appendix D.3.2.

4.6 Optimal Policy Structure for Single-Minded RS

We begin by presenting our main result, which characterizes the optimal recommendation policy in the single-minded setting—where the RS aims solely to maximize either user welfare or platform quality. Notably, the result shows that fairness, in the form of the uniform exposure policy UNIFORMBUTLAST, can emerge endogenously when content creation costs are convex and the platform prioritizes content quality ($\alpha = 0$).

Restatement of Theorem 4.1.2 (Optimal Policy in Boundary Regimes)**.** *Depending on the weight parameter α and the cost parameter β, the following holds:*

1. *If $\alpha = 1$ or $\beta \leq 1$, then* HARDMAX *is optimal.*

2. *If $\alpha = 0$ and $\beta < 1$, then* HARDMAX *is optimal.*

3. *If* $\alpha = 0$ *and* $\beta > 1$*, then* UNIFORMBUTLAST *is optimal.*

4. *If* $\alpha = 0$ *and* $\beta = 1$*, then any* $\boldsymbol{p} \in \Delta$ *with* $p_n = 0$ *is optimal.*

To understand this result, we analyze the RS's objective in the single-minded case as a function of the policy vector **p**, and reveal key structural properties using tools from majorization theory.

We first introduce a symmetric extension of the function $h(x, \mathbf{p}) = \sum_{i=1}^{n} a_i(x)p_i$, defined over the entire n-dimensional probability simplex Δ_n, not necessarily ordered. Let $O(x, \mathbf{p}) = \sum_{i=1}^{n} a_i(x)p_{(i)}$, where $p_{(i)}$ is the i-th largest component in **p**.

This construction allows us to study $O(x, \mathbf{p})^r$ as a symmetric function, which we relate to the concepts of majorization and Schur-convexity.

Definition 4.6.1 (Symmetric function)**.** A function $f : X^n \to \mathbb{R}$ is symmetric if $f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ for any permutation σ and any $x \in X^n$.

Definition 4.6.2 (Majorization)**.** Given $x, y \in \mathbb{R}^n$, we say x majorizes y ($x \succ y$) if:

$$\sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)} \quad \text{for all } k = 1, \ldots, n-1,$$
$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i,$$

where $x_{(i)}$ and $y_{(i)}$ are sorted in non-increasing order.

Definition 4.6.3 (Schur-convexity)**.** A symmetric function $f : \mathbb{R}^n \to \mathbb{R}$ is Schur-convex if $x \succ y$ implies $f(x) \geq f(y)$, and Schur-concave if $f(x) \leq f(y)$.

We use the following criterion to check Schur-convexity:

Lemma 4.6.4 (Schur-Ostrowski Criterion [102]). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be symmetric and continuously differentiable. Then f is Schur-convex (concave) if and only if:*

$$(x_i - x_j)\left(\frac{\partial f}{\partial x_i} - \frac{\partial f}{\partial x_j}\right) \geq (\leq)\ 0,$$

for all i, j.

Our main structural result is stated below.

Lemma 4.6.5 (Schur Properties of Objective). *For any $\boldsymbol{p} \in \Delta_n$, the function $\int_0^1 O(x, \boldsymbol{p})^r dx$ is:*

- *Schur-concave if $r \in [0, 1]$;*
- *Schur-convex if $r \leq 0$ or $r \geq 1$.*

Theorem 4.1.2 follows from applying Lemma 4.6.5 to the RS objective in the respective cases of α and β. A full proof is provided in Appendix D.4.2.

4.7 Optimal Policy Structure for Convex-minded RS

In this section, we characterize the optimal recommendation policy when the RS has a convex-minded objective, i.e., it balances user welfare and platform quality with a weight parameter $\alpha \in (0, 1)$. Let $G(\mathbf{p})$ denote the objective function associated with a given α as defined in OPT (see Theorem 4.1.1).

We restrict the domain of G to the ordered simplex Δ, as introduced in Section 7.2, to avoid confusion with its symmetric extension $O(x, \mathbf{p})$. From Theorem 4.1.1, we know that for $n = 2$, the optimal policy always satisfies $p_n = 0$, implying that the optimal policy in this case is HARDMAX. Therefore, we focus on the case where $n \geq 3$.

We now state our main result:

Restatement of Theorem 4.1.3 (Optimal Structure for Convex-Minded RS)**.** *For any* $\alpha \in [0, 1]$ *and* $\beta > 0$*, any optimal policy has the structure:*

$$p_1 \geq p_2 = p_3 = \ldots = p_{n-1} \geq p_n = 0.$$

4.7.1 Proof Sketch of Theorem 4.1.3

We prove this theorem by analyzing the geometry of the partial derivatives $\partial G(\mathbf{p})/\partial p_i$ for $i \in [n-1]$ and applying the KKT conditions to characterize the structure of optimal solutions.

4.7.1.1 Step 1: Quasiconvexity of the Gradient Sequence

Define the partial derivatives as:

$$d_i := \frac{\partial G(\mathbf{p})}{\partial p_i} = n\alpha \left(1 + \frac{1}{\beta}\right) \int_0^1 h(x, \mathbf{p})^{1/\beta} a_i(x) dx + \frac{1-\alpha}{\beta} \int_0^1 h(x, \mathbf{p})^{1/\beta - 1} a_i(x) dx, \tag{4.2}$$

for $i \in [n-1]$. Factoring out common factors, this can be written as

$$d_i = \int_0^1 \left[n\alpha \left(1 + \frac{1}{\beta} h(x, \mathbf{p})^{1/\beta} + \frac{1-\alpha}{\beta} h(x, \mathbf{p})^{1/\beta - 1}\right)\right] a_i(x) dx. \tag{4.3}$$

In fact, notice that for any general objective function of the form

$$H(\mathbf{p}) = e_1 \int_0^1 h(x, \mathbf{p})^{r_1} dx + e_2 \int_0^1 h(x, \mathbf{p})^{r_2} dx,$$

for $e_1, e_2, r_1, r_2 \in \mathbb{R}$, we can write its partial derivatives with respect to p_i's as:

$$\frac{\partial H(\mathbf{p})}{\partial p_i} = \int_0^1 q(x) a_i(x) dx, \tag{4.4}$$

for some $q(x)$ which involves terms n, α, β and $h(x, \mathbf{p})$. Importantly, $q(x)$ in the integrand remains the same across $i \in [n-1]$, which will play a key role in our analysis.

Looking forward, we will show that the sequence $(d_i)_{i \in [n-1]}$ exhibits quasiconveixty (along with some further properties), defined below:

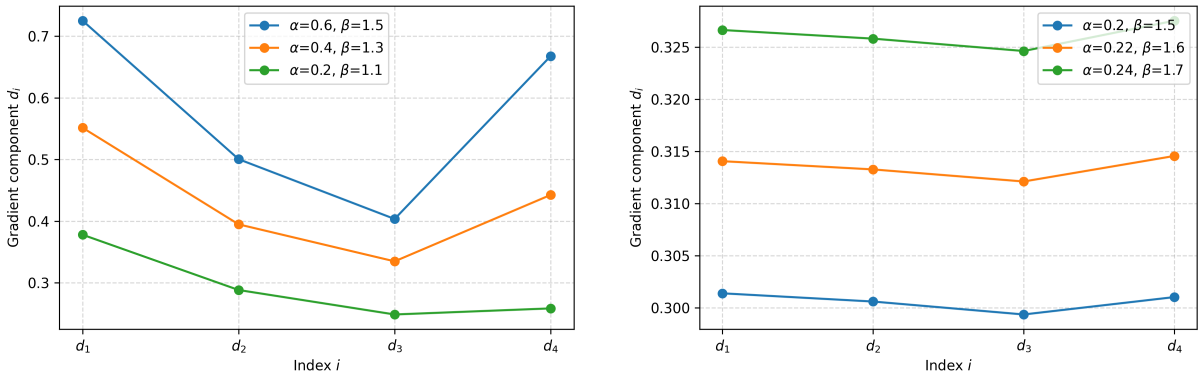
Definition 4.7.1 (Quasiconvexity)**.** A sequence $x_1, \ldots, x_n$ is *quasiconvex* if there exists $k \in [n]$ such that $x_1 \geq x_2 \geq \ldots \geq x_k \leq x_{k+1} \leq \ldots \leq x_n$. Similarly, a function $f : X \subseteq \mathbb{R} \to \mathbb{R}$ is quasiconvex if there exists $x^* \in \mathbb{R}$ such that $f(x) \geq f(y)$ for $x < y \leq x^*$ and $f(x) \leq f(y)$ for $x^* \leq x < y$.

In fact, we will prove a rather general statement such that if the partial derivatives of the objective function with respect to p_i's can be written as (4.4), and if the corresponding weight $q(x)$ behind each $a_i(x)$ is quasiconvex, then the resulting sequence of partial derivatives is highly structured. This is formally captured by the following key structural lemma:

Lemma 4.7.2 (Gradient Quasiconvexity and Structure)**.** *Given any general objective function $H(\boldsymbol{p})$, suppose its partial derivatives with respect to p_i's can be written as (4.4). If $q(x)$ in the integrand of (4.4) is quasiconvex on $x \in [0, 1]$, then the sequence $d_1, d_2, \ldots, d_{n-1}$ satisfies:*

1. *It is quasiconvex.*

2. *$d_i = d_{i+1}$ can only happen in one of the following cases: (i) if k is the slope-changing index where $p_{k-1} \geq p_k \leq p_{k+1}$, then $d_i = d_{i+1}$ only if $i = k-1$ or $i = k$, (ii) if the entire sequence d_i's for $i \in [n-1]$ is monotone decreasing, then $d_i = d_{i+1}$ only if $i = n-2$, and (iii) if the*



(a) d_i under HARDMAX policy

(b) d_i under UNIFORMBUTLAST policy

Figure 4.3: Behavior of d_i for $n = 5$. Since $p_5 = 0$, d_5 is omitted. Quasiconvexity of d_i follows from the quasiconvexity of $q(x)$ shown in Figure 4.4.

entire sequence d_i's for $i \in [n-1]$ is monotone increasing, then $d_i = d_{i+1}$ only if $i = 1$ or the entire sequence is monotone decreasing, in which case $d_{n-2} = d_{n-1}$.

3. *If $d_k = d_{k+1}$, then $d_k \neq d_{k-1}$, and vice versa.*

We prove Lemma 4.7.2 using tools from total positivity and the variation diminishing property. Essentially, totally positive operators ($a_i(x)$'s in our case) do not increase the number of sign changes, so the quasiconvexity in the operand $q(x)$ over x translates to the quasiconvexity of d_i's over i.

Finally, combined with the following lemma stating the quasiconvexity of $q(x)$ in our objective function $G(\mathbf{p})$, we conclude that d_i's has the desired properties.

Lemma 4.7.3. *The function*

$$q(x) = \left\lceil n\alpha \left(1 + \frac{1}{\beta} h(x, \boldsymbol{p})^{1/\beta} + \frac{1-\alpha}{\beta} h(x, \boldsymbol{p})^{1/\beta - 1} \right) \right\rceil$$

in (4.3) is quasiconvex on $x \in [0, 1]$.

Total Positivity and Variation Diminishing Properties

We briefly review key definitions:

Definition 4.7.4 (Total Positivity)**.** Let $K(x, y)$ be a real-valued function over ordered sets X and Y. It is totally positive of order r (TP$_r$) if for all $1 \leq k \leq r$ and increasing sequences $x_1 < \ldots < x_k$ in X, and $y_1 < \ldots < y_k$ in Y,

$$\det\left(K(x_i, y_j)\right)_{i,j=1}^{k} \geq 0.$$

If the inequality is strict, it is strongly totally positive of order r (STP$_r$).

The notion of total positivity is closely related to how much sign change occurs in a sequence or a function. The number of sign change is formally defined as follows:

Definition 4.7.5 (Number of Sign Changes)**.** For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, let $S^-(x)$ be the number of strict sign changes, ignoring zeroes. $S^+(x)$ allows arbitrary signs for zero entries, maximizing the number of sign changes.

The following notion of variation diminishing property formally defines the connection between the number of sign changes in a resulting function, which is calculated by a weighted summation over an original function given a two-dimensional weight kernel.

Definition 4.7.6 (Variation Diminishing Property)**.** Let $K(x, y)$ be an integrable kernel and $q(x)$ an integrable function. Define

$$Q(y) = \int K(x, y)q(x)dx.$$

Then K has the variation diminishing (VD) property if $S^-(Q) \leq S^-(q)$ and the strong variation diminishing (SVD) property if $S^+(Q) \leq S^-(q)$. Further, if the equation holds, then the sign patterns

should remain the same.

Essentially, the following seminal result asserts that any totally positive kernel cannot change the number of sign changes dramatically, allowing the variation diminishing property:

Theorem 4.7.7 ([90], Chapter 5)**.** *If $K(x, y)$ is TP_r, then it has VD for any q with $S^-(q) \leq r - 1$. If K is STP_r, then it has SVD for any q with $S^-(q) \leq r - 1$.*

To relate this to quasiconvexity, we use the following lemma:

Lemma 4.7.8 ([92])**.** *A function $g : I \to \mathbb{R}$ is quasiconvex iff for every $\lambda \in \mathbb{R}$, the function $f_\lambda(x) := g(x) - \lambda$ has at most two strict sign changes, with a sign pattern $(+, -, +)$ when the bound is tight.*

We also use the following result on generalized Vandermonde matrices:

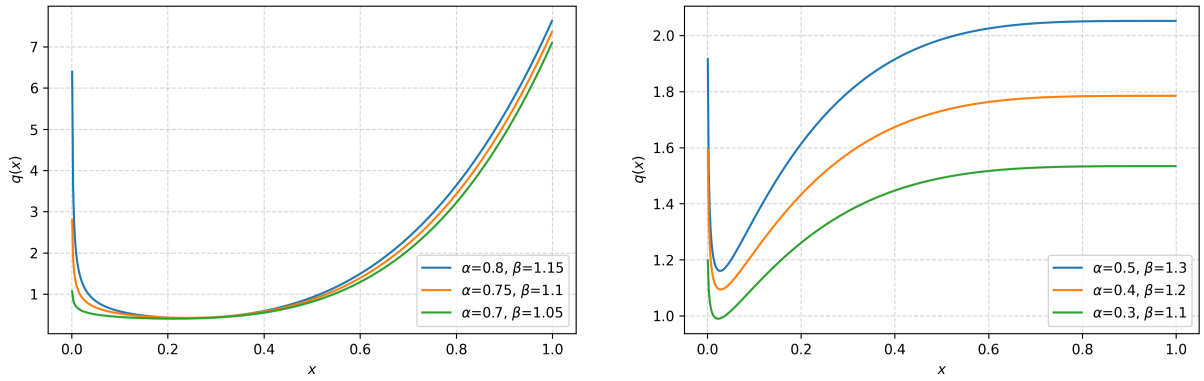
Theorem 4.7.9 ([91])**.** *Let $x_1 < \ldots < x_n$ be positive real numbers and $a_1 > \ldots > a_n$ be integers. Then,*

$$\det \left(x_i^{a_j}\right)_{i,j=1}^{n} > 0.$$

We defer the full proof of Lemma 4.7.2 and Lemma 4.7.3 to Appendix D.5.1. Intuitively, the quasiconvexity of $q(x)$ in Lemma 4.7.3 translates to the quasiconvexity of d_i's, thanks to the total positivity of our kernel function $a_i(x)$. This is illustrated in Figure 4.4.

4.7.1.2 Step 2: Applying KKT Conditions

With Lemma 4.7.2 in hand, we analyze the KKT conditions for the optimization problem OPT. These involve the derivatives d_i, the probabilities p_i, and associated dual variables. Using the structure guaranteed by Lemma 4.7.2, we show via a case analysis that any deviation from the structure in



(a) $q(x)$ under HARDMAX policy

(b) $q(x)$ under UNIFORMBUTLAST policy

Figure 4.4: Behavior of $q(x)$ for $n = 5$ (see Appendix D.5.3 for definition). The quasiconvexity of $q(x)$ transfers to d_i due to variation diminishing properties of the kernel defined by $a_i(x)$.

Theorem 4.1.3 would violate the KKT conditions. In particular, the transition in probability weights must occur only at indices 1 and n.

The full proof is provided in Appendix D.5.3.

4.7.2 Implication on High-dimensional Nonconvex Optimization Problem

Recall that Lemma 4.7.2 holds for any objective function whose partial derivatives are of the form (4.4), and its sequence is quasiconvex. This immediately implies an interesting corollary on a high-dimensional nonconvex optimization. Specifically, consider a monotone increasing nonnegative function $f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ that may possibly be nonconvex, a constant $C > 0$, and the following

constrained optimization problem:

$$\begin{aligned} (\text{GEN}) \quad & \underset{\mathbf{x}\in\mathbb{R}^n}{\text{maximize}} \quad f(\mathbf{x}) \\ & \text{subject to} \quad x_1 \geq x_2 \geq \ldots \geq x_n \geq 0 \\ & \qquad\qquad \sum_{i=1}^{n} x_i = C \end{aligned}$$

Then, by following the proof steps of Theorem 4.1.3, one can obtain the following general result.

Restatement of Corollary 4.1.4. *Consider the constrained optimization problem* GEN. *Suppose that the gradient of f with respect to each x_i can be written as the following:*

$$\frac{\partial f(\boldsymbol{p})}{\partial x_i} = \int_D w(y,\boldsymbol{p}) g_i(y) dy,$$

for a continuous bounded interval D, a function $g_i : D \to \mathbb{R}_{\geq 0}$ for $i \in [n]$, and another function $w(y,\boldsymbol{p}) : D \times \mathbb{R}^n \to \mathbb{R}$. Assume further that $g_i(y)$ is STP$_n$,[25]*,* i.e., *strongly totally positive of order n, and $w(y,\boldsymbol{p})$ is quasiconvex over $y \in D$,* i.e., *it decreases then increases, given any $\boldsymbol{p}$. Then, it follows that the optimum of f satisfies $x_1 \geq x_2 = \ldots = x_n$.*

We omit the proof as it follows from the exactly same proof structure of Theorem 4.1.3, equipped with required lemmas.

The implication of this structural characterization is significant, since once we know that $x_1 \geq x_2 = \ldots = x_n = x$, due to the monotonicity of f, it is immediate that $x_1 = C - (n-1)x$. Thus, the

[25] We refer to Definition 4.7.4 for more details.

optimization can be reduced to the following.

$$(\text{GEN'}) \quad \underset{x \in \mathbb{R}}{\text{maximize}} \quad f(C-(n-1)x, x, x, \ldots, x)$$
$$\text{subject to} \quad 0 \le x \le C/(n-1)$$

Notably, this is a single-dimensional optimization problem over x on a bounded interval, which typically admits efficient algorithms when f exhibits desirable properties such as smoothness or Lipschitzness.

4.8 Computing Optimal Policy

Now we know that any optimal policy satisfies $p_1 \ge p_2 = \cdots = p_{n-1} \ge p_n = 0$, we have $p_1 = p_2 = \cdots = p_{n-1} = \frac{1-p_1}{n-2}$ and $p_n = 0$, so the computation of the optimal policy reduces to find $p_1 \in \left[\frac{1}{n-1}, 1\right]$. Let us define $\mathcal{D} = [1/(n-1), 1]$, and write $h(x, p_1)$ to denote the corresponding $h(x, \mathbf{p})$ given the structural relation. Recall that $a_i(x) = \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}$, $h(x, p_1) = \sum_{i=1}^{n-1} a_i(x) p_i$, and then the objective function becomes

$$G(p_1) = \alpha n \int_0^1 h(x, p_1)^{1+1/\beta} dx + (1-\alpha) \int_0^1 h(x, p_1)^{1/\beta}\, dx,$$

where $h(x, p_1) = c_0(x) + c_1(x) p_1$ with

$$c_0(x) = \frac{1 - a_1(x) - a_n(x)}{n-2}, \qquad c_1(x) = a_1(x) - \frac{1 - a_1(x) - a_n(x)}{n-2}.$$

We will show that the resulting optimization over a single variable p_1 can be efficiently solved

using branch-and-bound algorithms, despite its nonconvexity. To this end, we first identify a lower bound and upper bound for our objective functions given an interval.

For an interval $I = [\ell, u] \subseteq \mathcal{D}$ define the lower and upper bounds on the *objective value* over I by

$$\begin{aligned} L(I) &:= \max\{G(\ell),\, G(u)\}, \\ U(I) &:= \alpha n \int_0^1 \max\{h(x,\ell)^{1+1/\beta}, h(x,u)^{1+1/\beta}\}\, dx \\ &\qquad + (1-\alpha) \int_0^1 \max\{h(x,\ell)^{1/\beta}, h(x,u)^{1/\beta}\}\, dx. \end{aligned}$$

We then first show that these constitute upper and lower bounds on the objective function.

Lemma 4.8.1. *For every $I = [\ell, u] \subseteq \mathcal{D}$,*

$$L(I) \;\le\; \sup_{p_1 \in I} G(p_1) \;\le\; U(I).$$

The above bounds will be crucial in our algorithm. Let us write $|I| := u - \ell$ to denote the length of the interval. The following lemma derives a Lipscthiz-style of bounds in the difference $U(I) - L(I)$.

Lemma 4.8.2. *For every interval $I = [\ell, u] \subseteq \mathcal{D}$,*

$$U(I) - L(I) \;\le\; C_1\,|I| \;+\; C_2\,|I|^{1/\beta},$$

where

$$C_1 = \alpha n\Big(1 + \tfrac{1}{\beta}\Big) \int_0^1 |c_1(x)|\, dx, \qquad C_2 = (1-\alpha) \int_0^1 |c_1(x)|^{1/\beta}\, dx.$$

Remark 4.8.3. Our algorithm and analysis will depend on parameters C_1 and C_2. In fact, instead of

computing the exact values, one can replace them with rough upper bounds. This would only incur a constant factor blowup in the running time. Formally, using $|c_1| \le a_1 + \frac{1-a_1-a_n}{n-2}$, we obtain

$$\int_0^1 |c_1(x)|\,dx \;\le\; \frac{1}{n} + \frac{1-2/n}{n-2} = \frac{2}{n}.$$

Note that by $(x+y)^\rho \le x^\rho + y^\rho$ when $\rho \in (0,1]$ and $a_1^{1/\beta}$ integrable,

$$\int_0^1 |c_1(x)|^{1/\beta}\,dx \;\le\; \frac{\beta}{\beta+n-1} + \frac{1}{(n-2)^{1/\beta}}.$$

Thus one may take

$$C_1' = 2\alpha\Big(1+\tfrac{1}{\beta}\Big), \qquad C_2' = (1-\alpha)\left(\frac{\beta}{\beta+n-1} + \frac{1}{(n-2)^{1/\beta}}\right),$$

and follow the analogous arguments.

Our main algorithm is presented in Algorithm 1. To analyze its correctness and time complexity, we first verify that the optimum is always bounded.

Lemma 4.8.4. *At every iteration: (i)* $L^\star \le \mathrm{OPT} := \max_{p_1 \in \mathcal{D}} G(p_1)$*; (ii)* $\max_{I \in \mathcal{A}} U(I) \ge \mathrm{OPT}$.

For ε-additive error, we ultimately require the following number of bisecting intervals.

Definition 4.8.5. Given $\varepsilon > 0$, define

$$\delta(\varepsilon) := \min\left\{\frac{\varepsilon}{C_1},\ \left(\frac{\varepsilon}{C_2}\right)^\beta\right\}.$$

Finally, we can show that the algorithm terminates in logarithmic over $1/\varepsilon$.

ALGORITHM 1: Active-Set Branch-and-Bound

Input: Tolerance $\varepsilon > 0$
Output: ε-optimal policy $\mathbf{p}^\star$
Initialize the active set $\mathcal{A} \leftarrow \{I_0\}$ with $I_0 = [1/(n-1), 1]$
For each $I \in \mathcal{A}$, compute $L(I), U(I)$
Set $L^\star \leftarrow \max_{I \in \mathcal{A}} L(I)$ and record endpoint $p^\star$ attaining $L^\star$
while $\max_{I \in \mathcal{A}} U(I) > L^\star + \varepsilon$ **do**
  Pick $I = [\ell, u] \in \mathcal{A}$ with the largest $U(I)$
  if $U(I) \le L^\star + \varepsilon$ **then**
    Remove I from $\mathcal{A}$ (prune)
  else
    Bisect I at midpoint $m = (\ell + u)/2$ into $I_1 = [\ell, m]$ and $I_2 = [m, u]$
    Compute $L(\cdot), U(\cdot)$ for I_1, I_2 and replace I by I_1, I_2 in $\mathcal{A}$
    Update $L^\star$ and $p^\star$ if any child's $L(\cdot)$ improves the incumbent
return $\boldsymbol{p}^\star = (p^\star, (1-p^\star)/(n-1), \ldots, (1-p^\star)/(n-1), 0)$

Theorem 4.8.6. *Consider a value oracle that exactly computes our objective function given a policy. Define*

$$C_1 = \alpha n\Big(1 + \tfrac{1}{\beta}\Big)\int_0^1 |c_1(x)|\,dx, \qquad C_2 = (1-\alpha)\int_0^1 |c_1(x)|^{1/\beta}\,dx, \qquad D = 1 - \frac{1}{n-1}.$$

The algorithm halts and returns $p^\star$ with $G(p^\star) \ge \mathrm{OPT} - \varepsilon$. Moreover, if an interval has undergone k bisections then its length is at most $D\,2^{-k}$, and it suffices that

$$k \;\ge\; \max\left\{\log_2 \frac{C_1 D}{\varepsilon}\,,\; \beta \log_2 \frac{C_2 D^{1/\beta}}{\varepsilon}\right\}$$

to guarantee that every active interval has length $\le \delta(\varepsilon)$. Further, the total number of nodes explored is $O\Big(\frac{C_1}{\varepsilon}\Big) \;+\; O\Big(\frac{C_2}{\varepsilon}\Big)^{\beta}$, i.e., *the overall running time is polynomial in $1/\varepsilon$.*

Approximating integrals. Note that our value oracle requires exactly calculating integrals, which necessarily entails numerical errors. That being said, if each integral is approximated with absolute

error at most δ, the returned policy $\mathbf{p}^\star$ satisfies

$$G(\mathbf{p}^\star) \;\geq\; \mathrm{OPT} - (\varepsilon + 2\delta).$$

Thus, setting $\delta \leq \varepsilon/4$ and running the same algorithm with target $\varepsilon/2$ would keep the overall error less than ε.

Thus, the natural question is whether we can efficiently compute the approximate integrals.

Lemma 4.8.7 (Approximate integrals)**.** *Fix $\beta > 0$ and $p \in \Delta$. Let $r \in \{1/\beta,\, 1 + 1/\beta\}$ and define $f_r(x) := h(x,p)^r$. Then $f_r : [0,1] \to [0,1]$ is monotone increasing*[26] *and for any integer $m \geq 1$,*

$$\left| \int_0^1 f_r(x)\,dx - \frac{1}{m}\sum_{j=1}^{m} f_r\!\left(\tfrac{j}{m}\right) \right| \;\leq\; \frac{f_r(1) - f_r(0)}{m} \;\leq\; \frac{1}{m}.$$

In particular, choosing $m = \lceil 1/\varepsilon \rceil$ yields a deterministic additive-ε approximation to $\int_0^1 h(x,\boldsymbol{p})^r dx$ using m evaluations of $h(x,\boldsymbol{p})^r$.[27]

Noting that computation of $h(x,\mathbf{p})^r = h(x,p_1)^r = (c_0(x) + c_1(x)p_1)^r$ with r as a function of β only requires standard arithmetic operations running in polynomial over n given a fixed β, the overall running time is polynomial in n and $1/\varepsilon$.

4.9 Beyond RS: General Objective Functions

In hindsight, we notice that the proof of Lemma 4.7.2 only requires the platform's objective function to have certain property, where our objective function $G(\mathbf{p})$ as stated in Theorem 4.1.1 turns

[26] By Lemma D.5.1, $h(\cdot,\mathbf{p})$ is strictly increasing for any nontrivial $\mathbf{p}$, hence $f_r(\cdot)$ is increasing for $r > 0$.

[27] Note that for multiplicative approximation, one can stop when $\max_I U(I) \leq \frac{1}{1-\eta} L^\star$ yields $G(\mathbf{p}^\star) \geq (1-\eta)\mathrm{OPT}$.

out to meet this condition. This inspires us to derive a significantly more general condition under which our analysis on the characterization of optimal policies carries over. Interestingly, we observe that our analysis holds for any posynomial objective function with nonnegative exponents.[28]

Restatement of Theorem 4.1.5 (Posynomial objective function)**.** *Given* $k_m > k_{m-1} > \ldots > k_1$*, consider the following generalized objective function at symmetric MNE* $\bar{\mu}$ *given* $\boldsymbol{p}$*:*

$$\mathcal{G}(\bar{\mu}, \boldsymbol{p}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{q_i \sim \mu_i} \left[e_m q_i^{k_m} + e_{m-1} q_i^{k_{m-1}} + \ldots + e_1 q_i^{k_1} \right],$$

where $q_i \sim \mu_i$ *for a symmetric strategy profile* μ_i *in* $\bar{\mu}$*. For any* $\beta > 0$*, suppose* e_j*'s and* k_j*'s satisfy either of the following property:*

1. *There exists* $j^* \in [n-1]$ *such that* $e_j(k_j - \beta) \leq 0$ *for* $j \leq j^*$ *and* $e_j(k_j - \beta) \geq 0$ *for* $j > j^*$*.*
2. *All the elements in the sequence* $e_j(k_j - \beta)$ *for* $j \in [n-1]$ *is nonnegative or nonpositive.*

Then, the Lemma 4.7.2 holds, i.e.*, any optimal policy has the same structure as Theorem 4.1.3,* i.e.*,*

$$p_1 \geq p_2 = p_3 = \ldots = p_{n-1} \geq p_n = 0.$$

As a special case, if e_j*'s are all nonnegative, this condition follows.*

The proof is based on the observation that given the posynomial function, the weights (except $a_i(x)$) in the integrand of its partial derivative with respect to p_i's exhibits quasiconvexity over $x \in [0, 1]$, which is precisely the condition for Lemma 4.7.2. We defer the proof to Appendix D.7.

[28] Posynomial refers to a function which is a positive combination over monomials. With a single variable, it can be viewed as a polynomial with nonnegative coefficients whose exponent can possibly be noninteger.

Social welfare. Our characterization under the general objective function in Theorem 4.1.5 implies several interesting applications. For instance, let us consider *social welfare* of the contest, which is the summation of the RS's utility and the producers' utilities. Let $\mathcal{V}(\bar{\mu}, \mathbf{p})$ be the platform's utility function at symmetric MNE induced by $\mathbf{p}$, *e.g.,* average quality or a function of it. Then, at a symmetric MNE $\bar{\mu} = (\mu, \dots, \mu)$, the social welfare can be written as:

$$\mathcal{S}(\bar{\mu}, \mathbf{p}) = \mathcal{W}(\bar{\mu}, \mathbf{p}) + \mathcal{V}(\bar{\mu}, \mathbf{p}) + \sum_{i \in [n]} \mathcal{U}_i(\mu, \bar{\mu}_{-i}, \mathbf{p}). \tag{4.5}$$

Then, Theorem 4.1.5 immediately implies the following corollary:

Corollary 4.9.1. *Assume the platform's utility function is given by*

$$\mathcal{V}(\bar{\mu}, \boldsymbol{p}) = \mathbb{E}_{q \in \mu} \left[e_m q^m + \dots + e_1 q \right],$$

where $e_j \geq 0$ for all $j \in [m]$. Suppose the RS aims to maximize the social welfare as written in (4.5). Then, the optimal policy has the structure $p_1 \geq p_2 = \dots = p_{n-1} \geq p_n = 0$ at the symmetric MNE.

Inverse S-shaped function. On the other hand, the RS might be interested in a *superstar* outcome of the contest where it realizes a concave utility up to certain quality threshold, and then it realizes a convex utility of increasing marginals.

For instance, assume the convex cost $\beta = 2$, and the following utility function:

$$\mathbb{E}_{q \sim \mu} \left[2q^3 - 3q^2 + 2q \right].$$

One can easily verify that this is monotone in q, and also satisfies the condition in Theorem 4.1.5.

Notably, this function is inversely s-shape, *i.e.,* strongly concave-then-convex over q. In words, the RS's utility does not significantly change up to certain threshold quality level, but then it dramatically increases afterwards. This is in a similar vein to the S-shaped function in prospect theory [95] in behavioral economics, but in an inverse manner as standard S-shaped function requires the utility function to be convex-then-concave. In practice, many research contests that requires innovative outcome would satisfy this property. Further, one can imagine of a economies of scale, *e.g.,* if the outcome of the contest exceeds certain threshold, then the company (contest designer) can monopolize the market to bring increasingly larger marginal revenues.

Order statistics. Finally, one might be interested in the highest quality contents among the producers, *i.e.,* $\mathcal{V}(\bar{\mu}, \mathbf{p}) = \mathbb{E}\left[q_{(1)}\right]$ where $q_{(1)}$ is the largest order statistics among $q_1, \ldots, q_n$. The following corollary suggests that the structure of the optimal policies remains the same for such objective.

Corollary 4.9.2. *Assume the platform aims to maximize the expected value of the highest quality outcome,* i.e., $\mathbb{E}_{q_i \sim \mu_i, i \in [n]}\left[q_{(1)}\right]$, *where* $q_{(1)}$ *denotes the largest order statistics. Then, the optimal policy has the structure* $p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$ *at the symmetric MNE.*

Exponential utility. Even more extreme version of the inverse S-shaped function can be represented as an exponential utility where the utility function is written as $\mathbb{E}_{q \sim \mu}\left[e^{\lambda q}\right]$. Notice that Theorem 4.1.5 accommodates the exponential utility as shown in the following.

Corollary 4.9.3. *Assume the platform aims to maximize the exponential utility,* i.e., $\mathbb{E}_{q_i \sim \mu_i, i \in [n]}\left[e^{\lambda q_i}\right]$, *for* $\lambda > 0$. *Then, the optimal policy has the structure* $p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$ *at the symmetric MNE. Further, the same structure holds if the platform aims to maximize a convex combination of exponential utilities* $\mathbb{E}_{q_i \sim \mu_i, i \in [n]}\left[\sum_{j=1}^{m} e^{\lambda_j q_i}\right]$.

This follows from uniformly approximating the exponential function with its Taylor series, the coefficients of which satisfy the condition of Theorem 4.1.5.

Chapter 5: Replication-proof Bandit Mechanism Design with Bayesian Agents

This chapter presents our work [23], which is a Bayesian extension of [22].

5.1 Introduction

Multi-armed bandit (MAB) algorithms are an important paradigm for interactive learning that focuses on quantifying the trade-off between exploration and exploitation [103], with various real-world applications such as recommender systems [104, 105], dynamic pricing by [106, 107], and clinical trials by [108] to name a few. They are used to model various practical applications where there is a set of actions (arms) to be selected (pulled) over a collection of rounds. Pulling an arm yields a reward sampled from its distribution which is generally different from the other arms' distributions. The objective is to maximize the sum of rewards accumulated through the rounds or equivalently minimize the *regret* with respect to the optimal clairvoyant arm choice.

In deploying MAB algorithms in real-world applications, however, the rewards are often not coming from stochastic distributions, but from a collection of rational individuals (agents) who seek to maximize their own payoff. Consequently, there has been a surge of interest in bandit problem from the mechanism design perspective, which studies MAB algorithms that induce a good outcome with respect to agents' strategic behavior, while maintaining its efficiency. Examples of such settings

include two-sided matching markets, where agents learn their preferences through sequential interactions [109, 110], and learning in environments where agents can strategically modify the realized rewards of their arms [111, 112].

Recently, [22] considered a problem where agents register arms to a platform that uses a MAB algorithm to minimize regret. Each agent has an original set of arms, and the mean reward of each arm is known to him a priori, *i.e.,* agents are *fully-informed*. Each agent receives a *fixed constant portion* of the reward whenever its arm is pulled, and tries to maximize cumulative reward. Importantly, each agent can potentially register the same arm *more than once* (replication) at free cost, which possibly ruins the exploration-exploitation balance of the bandit algorithm. Indeed, [22] show that "ordinary" algorithms such as upper-confidence-bound (UCB) induce rational agents to replicate infinitely many times, and thus the algorithm suffers linear regret. To address this problem, [22] propose a hierarchical algorithm with two phases, and prove its truthfulness such that each agent's dominant strategy is not to replicate.

Although the introduction of replication-proof mechanism is remarkable, the assumption that each agent is fully informed about its own arms is theoretically strong and practically limited. Intuitively, if an agent knows all of its own arms' mean rewards, there exists no reason for the agent to *consider suboptimal arms*. Indeed, this assumption significantly simplifies the equilibrium analysis so that the problem simply reduces to the case under which each agent only has a *single arm*. Then, it is immediate to check that any bandit algorithm is replication-proof under a single-agent setting. Moreover, in practice, it is more reasonable to assume that each agent only has partial information regarding its own arms. For instance, in online content platforms, each content provider usually does not know the quality of each content that has not been posted yet, but only can guess its outcome, based on the data possibly acquired from the previous contents he/she may have.

In this context, we study a *Bayesian extension*[1] of the replication-proof bandit mechanism design problem, where each agent only knows such a distribution from which its own arm's mean reward is sampled.[2] Given the prior distribution, each agent computes its ex-ante payoff by taking expectation over all the possible realization of its own arms, and commits to a strategy that maximizes the payoff. Importantly, we aim to obtain a *dominant strategy incentive-compatible* (DSIC) mechanism such that each agent has no incentive to deviate from its truthful strategy, *i.e.,* submitting all the original arms without any replication, regardless of the others' strategies.

While our extension itself is intuitive and simple, this brings *significant challenges* in analyzing an algorithm's equilibrium compared to the fully-informed agent setting. Notably, even in a single-agent case, it is not trivial to obtain a replication-proof algorithm. This is in stark contrast to the fully-informed setting under which any bandit algorithm immediately satisfies replication-proofness. The question of designing replication-proofness in the single-agent setting spawns an intriguing question of *comparing expected regret of multiple bandit instances*, which will be elaborated shortly.

Outline of the paper. In the following subsection, we present our main contributions. Due to the page limit, we defer our discussion on related works to Appendix A. Then, we introduce the formal problem setup and preliminaries, and the failure of existing algorithms and its implication. Next we develop our main intuition on constructing replication-proof algorithm by investigating the single-agent setting. Correspondingly, we extend this result and provide our main result on replication-proof algorithm with sublinear regret. Practical motivations, discussion on an extension to asymmetric setting, details on the failure of H-UCB, and all the proofs are deferred to the appendix. Due to the page limit, all the

[1] We remark that our notion of Bayesian differs from that of Bayesian mechanism in the literature. While the Bayesian mechanism computes the equilibrium by considering ex-interim payoff of the players in the incomplete information setting, *i.e.,* ex-interim incentive-compatible, our mechanism is DSIC in an ex-post manner..

[2] We refer to Appendix B for more discussion on the practical motivation of our problem setup.

appendices can be found in the full paper [113].

5.1.1 Contributions and Techniques

Overall, our main results for replication-proof algorithms are presented in a step-by-step manner. We start with investigating the previous algorithm H-UCB suggested by [22], which mainly builds upon standard UCB along with the idea of hierarchical structure. Unlike from the fully-informed setting under which any bandit algorithm is replication-proof in the single-agent setting, we reveal that the standard UCB algorithm fails to be replication-proof even with a single-agent, regardless of the exploration parameter f, due to its deterministic (up to tie-break) and adaptive nature. This phenomenon carries over to the multi-agent setting, so that UCB with hierarchical structure (H-UCB) fails to be replication-proof, regardless of the number of agents.

To tackle this challenge, we first thoroughly investigate the single-agent setting, and characterize necessary and sufficient conditions for an algorithm to be replication-proof. Let us provide a simple example to explain the main technical challenges one would face in analyzing equilibriums.

Example 5.1.1. Consider a single agent with two Bernoulli arms, say arm a and b. The prior distribution is given by $Bern(0.5)$, *i.e.,* the expectation of both the arms' distributions are sampled from the prior distribution $Bern(0.5)$. The agent receives a constant portion, say 40% of the realized reward, whenever their arms are pulled. There exist four cases in terms of the pair of realized mean rewards (μ_a, μ_b): (i) $(1, 1)$, (ii) $(1, 0)$, (iii) $(0, 1)$, and (iv) $(0, 0)$, all of which occurs with probability 0.25. Due to the nature of Bayesian agents, the agent cannot observe which arm has larger mean rewards, and suppose that the agent somehow commits to a strategy that only replicates the arm a, say the replicated arm a'. For cases (i) and (iv), the replicated arm has no effect since all the arms have the same mean

rewards, *i.e.,* pulling any arm yields the same expected reward, so let's focus on case (ii) and (iii). In case (ii), the resulting bandit instance from the agent's strategy will be $\mathcal{I}_A : (\mu_a, \mu_{a'}, \mu_b) = (1, 1, 0)$ whereas in case (iii), it will be $\mathcal{I}_B : (0, 0, 1)$ Hence, compared to the strategy with no replication, the agent's gain from replicating arm a is equal to the gain from the average of expected payoff in case (ii) and (iii), compared to the case in which the algorithm runs with bandit instance $\mathcal{I}_O : (1, 0)$ that does not have any replicated arm. Hence, if the agent has no incentive to replicate arm a once, the expected regret under $\mathcal{I}_O$ should be at most the average of that under $\mathcal{I}_A$ and $\mathcal{I}_B$.

Even under simplistic and standard algorithms such as UCB, it is not straightforward to answer the above question. Intuitively, instance $\mathcal{I}_A$ will incur smaller regret as the number of optimal arm increases, but the instance $\mathcal{I}_B$ will have larger regret due to more suboptimal arms with mean reward 0.[3] Thus, the question is whether the loss from $\mathcal{I}_B$ is larger than the gain from $\mathcal{I}_A$ in average, compared to $\mathcal{I}_O$.

More formally, we can refine the above observation as the following fundamental question in comparing the expected regret of multiple bandit instances.

Given a bandit instance, consider an adversary who uniformly randomly selects an arm and replicates it. Which bandit algorithms guarantee that the regret does not increase against the adversary?

We affirmatively answer this question by proving that, there exists an algorithm that satisfies the above questions for *any* bandit instance under some mild assumptions. To this end, we introduce a notion of *random-permutation regret*, which captures the expected regret of an agent's strategy with

[3]This can be shown using a coupling-like argument between random reward tapes of the arms, combined with a regret decomposition lemma. See [22] for more details.

replication without observing the ex-post realization. Then, we show that if an algorithm satisfies (i) *truthfulness under random permutation* (TRP) such that the random-permutation regret does not increase with respect to replication, and (ii) permutation invariance such that the algorithm's choice does not depend on the arm indices, then the algorithm is replication-proof. More surprisingly, we prove that these conditions are indeed *necessary* for any algorithm to be replication-proof.[4] In fact, the failure of UCB comes from violating TRP. Meanwhile, we prove that exploration-then-commit (henceforth ETC) satisfies these properties, thus it is replication-proof in the single-agent setting.

For the multi-agent setting, we combine several techniques upon the observations above. First, we borrow the hierarchical structure of [22], and analyze hierarchical algorithm with each phase running ETC, which effectively decomposes each agent's payoff from each other, thanks to its non-adaptive nature. It remains as a major open problem whether it is necessary for the algorithm to have non-adaptive structure in order to satisfy TRP. In addition, we introduce a novel restarting round to re-initialize the intra-agent statistics. This largely simplifies the analysis of the selected agent in the exploitation rounds in the first phase, since the second phase simply reduces to the single-agent case with the highest empirical average reward. We cement our result by proving that our algorithm has sublinear regret.

5.2 Model

We consider a simultaneous game between a set of agents $[n] = \{1, 2, \ldots, n\}$ and a principal. Each agent $i \in \mathcal{N}$ has a set of original arms $\mathcal{O}_i = \{o_{i,1}, \ldots, o_{i,l_i}\}$ given a constant $l_i \in \mathbb{N}$. Each agent i is equipped with a cumulative distribution F_i from which each original arm $o_{i,k}$'s mean reward $\mu(o_{i,k})$ is sampled for $k \in [l_i]$. Each arm $o_{i,k}$ further follows a reward distribution $G_{i,k}$ with the mean reward

[4] We refer to Section 5.2 for precise definitions.

sampled as noted above. We use f_i and $g_{i,k}$ to denote the density function of F_i and $G_{i,k}$,[5] and we assume that all the distributions are supported on $[0, 1]$. We further assume that F_i is independent from F_j for $i \neq j \in [n]$. For instance, F_i might be a uniform distribution over $[0, 1]$, and each arm $a \in \mathcal{O}_i$ has a Bernoulli reward distribution with the mean sampled from F_i. We often deal with a *single agent* setting with $n = 1$.

Bayesian agents. Importantly, we consider a *Bayesian scenario* in which each agent is only aware of its own distribution F_i,[6] but not the exact realization of μ_a for $a \in \mathcal{O}_i$. We note that such assumption is indeed practical since content providers (agents) in contents platform typically only have a partial information on how much their contents may attract the users, *e.g.,* based on his/her previous contents and outcomes therein. This is in stark contrast to the setting of [22] in which they assume *fully-informed* agents such that any strategic agent does not have any incentive to register the arms except the best one reducing the problem into the single-arm case, which is not practical in many real-world scenarios. We highlight that the equilibrium analysis in the fully-informed setting simply boils down to the case under which each agent only has a single original arm and decides how much to replicate that unique arm, whereas our extension significantly complicates the analysis of equilibrium, as will be elaborated shortly.

Agent's strategy. Given the prior distribution F_i, the set of original $O_{i,k}$ and the principal's algorithm $\mathfrak{A}$, each agent i decides how many times to replicate arm k, including 0 of no replication, for each $k \in [l_i]$. Precisely, agent i commits to a set $\mathcal{S}_i = \{s_{i,k}^{(c)} : c \in [c_{i,k}], k \in [l_i]\}$ in a strategic manner, where $c_{i,k}$ denotes the total number of replica for arm $o_{i,k}$. The replicated arm $s_{i,k}^{(c)}$ has the same mean reward to the original arm, *i.e.,* $\mu(s_{i,k}^{(c)}) = \mu(o_{i,k})$ for any $i \in [n]$ and $k \in [l_i]$. We often use $s_{i,k}^{(0)}$ to denote $o_{i,k}$

[5] We assume the existence of density function for ease of exposition, but all the results will easily be generalized.
[6] Discussion on the generalization to the asymmetric setting is presented in Appendix C.

for simplicity. Define $\mathcal{S} = (\mathcal{S}_1, \ldots, \mathcal{S}_n)$ and $\mathcal{O} = (\mathcal{O}_1, \ldots, \mathcal{O}_n)$. Given all the agents' strategies $\mathcal{S}$, the bandit algorithm $\mathfrak{A}$ runs and corresponding rewards are realized.

Mechainsm procedure. The overall mechanism proceeds as follows: (i) The principal commits to a bandit algorithm (mechanism), (ii) Each agent $i \in \mathcal{N}$ decides an action $\mathcal{S}_i$ to register given $\mathfrak{A}$ to maximize own payoff, (iii) The mean rewards μ_a for each $a \in \mathcal{O}$ are sampled, (iv) The bandit algorithm $\mathfrak{A}$ runs and rewards are realized, (v) The principal and the agent realize their own utility.[7]

Agent's payoff. Once an arm is selected by $\mathfrak{A}$, a fixed portion $\alpha \in (0, 1)$ of the reward is paid to the agent who registers the arm. We write $\vec{\mu}_i$ to denote $(\mu_{i,1}, \mu_{i,2}, \ldots, \mu_{i,l_i})$, *i.e.,* the vector of agent i's (realized) mean rewards of the arms. Let $E_i(\vec{\mu}_i)$ be an event that the mean reward of agent i's arms are realized to be $\vec{\mu}_i$. Recall that $E_i(\cdot)$ are mutually independent for for $i \in [n]$. We often abuse E_i to denote $E_i(\vec{\mu}_i)$ if the realization $\vec{\mu}_i$ is clear from the context. We also write $\mathcal{S}_{-i}$ to denote $(\mathcal{S}_1, \ldots, \mathcal{S}_{i-1}, \mathcal{S}_{i+1}, \ldots, \mathcal{S}_n)$, and $E(\vec{\mu})$ for $\vec{\mu} = (\vec{\mu}_1, \vec{\mu}_2, \ldots, \vec{\mu}_n)$ to denote the event such that $\cap_{i \in [n]} E_i(\vec{\mu}_i)$.

We further introduce some notations given the principal's algorithm $\mathfrak{A}$. Conditioned on the $E_i(\vec{\mu}_i)$, agent i's *ex-post payoff* of selecting strategy $\mathcal{S}_i$ given the others' strategies is defined as

$$U_i(\mathcal{S}_i, \mathcal{S}_{-i}; E(\vec{\mu})) = \alpha \sum_{t=1}^{T} R_t \mathbf{1}\left\{ I^{(t)} \in \mathcal{S}_i \right\},$$

where R_t refers to the reward at round t and I_t denotes the arm selected at round t.

Given the ex-post utility, each agent computes its *ex-ante payoff* by marginalizing over the real-

[7]Note that if all the agents do not strategize and simply register $\mathcal{O}_i$ in step (ii), our problem reduces to the standard stochastic multi-armed bandit problem.

ization of its own arms' mean rewards, defined as follows

$$u_i(\mathcal{S}_i, \mathcal{S}_{-i}) = \mathbb{E}\left[\, U_i(\mathcal{S}_i, \mathcal{S}_{-i}; E(\vec{\mu}))\,\right],$$

where the randomness in the expectation comes from $E(\vec{\mu})$, i.e. the realization of the mean rewards, and possibly from the random bits of the algorithm $\mathfrak{A}$. In words, each agent is rational *expected utility maximizer* who maximizes its expected payoff.

Information and desiderata. We now describe the information structure of the mechanism, and how each agent decides its own strategy based on its ex-ante payoff. We first assume that the principal does not have any information regarding the agents' distributions nor the number of original arms, but can observe by whom each arm is registered. This is often called *prior-independent mechanism* in the mechanism design, *c.f.,* [114]. In addition, we assume that the agent does not have any information regarding the others. Instead, our objective is to construct a *incentive-compatible* mechanism such that each agent's *dominant strategy* is to register the arms without any replication, without regard to what others may have proposed. To this end, we first define a dominant strategy as follows.

Definition 5.2.1 (Dominant Strategy)**.** Agent i's strategy $\mathcal{S}_i$ is dominant strategy if $u_i(\mathcal{S}_i \cup \mathcal{S}_{-i}) \geq u_i(\mathcal{S}_i' \cup \mathcal{S}_{-i})$ for any $\mathcal{S}_i'$ and $\mathcal{S}_{-i}$.

Correspondingly, we introduce the following notion of *dominant strategy equilibrium*.

Definition 5.2.2 (Equilibrium)**.** An algorithm $\mathfrak{A}$ has dominant strategy equilibrium (DSE) if there exists a set of agent strategies $\mathcal{S} = \{\mathcal{S}_i\}_{i\in[n]}$ such that $\mathcal{S}_i$ is dominant strategy for any agent $i \in [n]$.

As per the definition, the total ordering of the agent's payoff with respect to its own strategy should remain the same regardless of the others' strategies $\mathcal{S}_{-i}$. Thus, we often omit the parameter

$\mathcal{S}_{-i}$ in denoting agent i's payoff, *e.g.,* $u_i(\mathcal{S}_i)$ instead of $u_i(\mathcal{S}_i, \mathcal{S}_{-i})$. Note, however, that the realization of the other agents' mean rewards may affect the actual quantity of $u_i(\mathcal{S}_i)$, although it does not affect which strategy agent i will eventually commit to. Since we are interested only in how each agent will "strategize", it does not lose any generality.

We finally introduce the following notion of truthfulness.

Definition 5.2.3 (Replication-proof)**.** An algorithm $\mathfrak{A}$ is *replication-proof* if it has a dominant strategy equilibrium $\{\mathcal{O}_i\}_{i\in[n]}$ for any $T \in \mathbb{N}$.

Replication-proofness is desirable as the agents are not motivated to replicate their own arms, thereby reducing the number of arms that a bandit algorithm faces. Since the regret upper bound of a standard bandit algorithm usually depends on the number of arms, maintaining a smaller number of suboptimal arms in the system affects the eventual learning efficiency of the algorithm. We often say that an algorithm is *incentive-compatible* or *truthful* to denote the replication-proofness [115].

On top of designing *prior-independent incentive-compatible* mechanism, we are also interested in a learning efficiency, *i.e.,* learning the reward distribution in a sample-efficient manner. In the literature, this is usually described as cumulative regret, defined as follows.

Definition 5.2.4 (Regret)**.** Given time horizon T, strategy $\mathcal{S}$ and the event $E(\vec{\mu})$, the ex-post regret of algorithm $\mathfrak{A}$ is

$$\textsc{Reg}(\mathfrak{A}, T; \mathcal{S}, E(\vec{\mu})) = \left(\max_{a\in\cup_{i\in\mathcal{N}}\mathcal{O}_i} \sum_{t=1}^{T} \mu(a) \right) - \sum_{t=1}^{T} \mu(I_t),$$

where I_t denotes the arm selected by $\mathfrak{A}$ at round t.

We emphasize that the optimal benchmark is selected from $\cup_{i\in\mathcal{N}}\mathcal{O}_i$, not from the registered set of arms $\cup_{i\in\mathcal{S}}\mathcal{S}_i$. This benchmark is indeed important since the principal should incentivize the agents

to decide an action that contains own best arm to maintain larger utility. For example, if an algorithm motivates an agent to not register all the arms, then there might be a constant probability that the omitted arm has the largest mean reward, which essentially yields sublinear regret with respect to the regret defined above. More formally, our notion of regret implicitly implies that any sublinear regret algorithm incentivizes all the agents to register all their original arms at least once. We also remark that although the arms' mean rewards are sampled from prior distributions, which is often referred to Bayesian bandits (*e.g.,* see Ch.3 of [103]), we deal with ex-post regret, *e.g.,* as does in [116] which analyzes the ex-post regret of Thompson Sampling in Bayesian bandits. Obviously, ex-post regret upper bound *implies* Bayesian regret upper bound by simply taking an expectation.

5.3 Failure of Existing Algorithms

Before getting into our main results, the natural question one might ask is to analyze whether the existing algorithms by [22] work. They mainly propose hierarchical UCB (H-UCB), which is a variant of the standard UCB by [32]. The formal pseudocode of both UCB and H-UCB are deferred to the appendix. While we provide the analysis with specific exploration parameter of $\ln t/n_a$, we note that the result easily carries over to arbitrary exploration parameters $f(t, n_a)$, once f is increasing over t and decreasing over n_a. We omit further details as it is beyond our interest.

Mainly, H-UCB consists of two phases: (i) in the first phase, it runs UCB1 by considering each agent as a single arm and selects one agent at each round, and (ii) in the second phase, it runs UCB within the selected agent. Precisely, it maintains two types of statistics: (i) agent-wise empirical average reward $\hat{\mu}_i$, number of pulls per agent n_i, and (ii) arm-wise empirical average reward $\hat{\mu}_{i,a}$ and number of pulls per arm $n_{i,a}$.

Theorem 5.3.1. *There exists a problem instance such that UCB1 is not replication-proof in the single-agent setting, and such that H-UCB is not replication-proof for any number of agents.*

The proof for the single-agent setting relies on a construction of bad problem instance which exploits the deterministic and adaptively exploring nature UCB. To briefly explain the proof, we consider an agent with two Bernoulli arms whose prior distribution is supported on $\{0, 1\}$, where it samples mean 0 and 1 equally likely. Let's focus on a realization in which one arm has mean 0, and the other has 1. Observe that replicating one arm essentially yields one of the following two bandit instances: (i) arms with mean rewards $(1, 1, 0)$ or (ii) arms with mean rewards $(1, 0, 0)$. Here, we write (x, y, z) to denote bandit instance with three arms mean rewards $1, 1, 0$. Note that, intuitively, (i) yields a lower expected regret, but higher in (ii), compared to the original instance with $(0, 1)$. Thus, our analysis mainly reduces to show that the loss from (ii) is larger than the gains from (i). This can be made indeed true since after observing the dynamics of instance (ii), we can essentially set the time horizon to be exactly right *before* the two arms with mean rewards zero are *consecutively* chosen. In this way, the loss from (ii) can be made smaller while the gains from (i) is moderately large, which implies the dominance of replicating strategy. The proof for the multi-agent setting essentially is based on the fact that its Phase 2 does not guarantee replication-proofness within the chosen agent due to the single-agent argument, along with carefully chosen problem parameters to propagate this phenomenon to Phase 1 so as to increase the agent's expected utility under replication.

5.4 Warm-up: Single-agent Setting

As noted thus far, UCB is not replication-proof even in the single agent setting. Indeed, as per Theorem 5.3.1, it is not obvious to verify which algorithm would satisfy replication-proofness even in

the single-agent setting. In this section, we step towards in constructing replication-proof algorithm, by carefully investigating the single-agent setting first. We present a set of conditions which is sufficient for an algorithm to be replication-proof in the single-agent setting, and then, as an example, will prove that exploration-then-commit (ETC) will satisfy these properties. As we restrict our attention to single-agent setting in this section, we often omit the agent index i in the notations.

To this end, we first introduce several notations. For any natural number l, let $\mathcal{P}_l$ be a set of all possible permutations $\sigma : [l] \mapsto [l]$. Now we introduce a dictionary form of bandit instance $\mathcal{I}$. Given the realization of the arms' mean rewards, suppose that the set of mean rewards constitute a sorted sequence $\{1 \geq \mu_1 > \mu_2 > \ldots > \mu_l > 0\}$, where some arms may share the same mean reward. For each μ_a for $a \in [l]$, we count the number of arms which have mean reward μ_a, and denote the count by c_a. Then, our bandit instance $\mathcal{I}$ can essentially presented as tuples of the mappings $\bigtimes_{a\in[l]}(\mu_a : c_a)$, which we say dictionary form of the bandit instance $\mathcal{I}$.

Suppose that we have a standard multi-armed bandit instance $\mathcal{I}$ with dictionary form $(\mu_a : c_a)_{a\in[l]}$. Then, we define a *permuted bandit instance* $\mathcal{I}_\sigma$ to be a bandit instance with dictionary form of $\bigtimes_{a\in[l]}(\mu_a : c_{\sigma(a)})$, ***i.e.,*** arms with mean reward μ_i appears $c_{\sigma(a)}$ times. For instance, given the bandit instance $\mathcal{I} = (0.5 : 1, 0.7 : 2)$, consider a permutation σ such that $\sigma(1) = 2$ and $\sigma(2) = 1$ correspondingly. Then, one can easily verify that $\mathcal{I}_\sigma = (0.5 : 2, 0.7 : 1)$.

Now, we introduce the random permutation regret.

Definition 5.4.1 (Random permutation regret)**.** Given a single-agent bandit instance $\mathcal{I}$ with l arms, we

define an algorithm $\mathfrak{A}$'s *random permutation regret* (RP-Regret) as follows.

$$\begin{aligned}\text{RP-REG}(\mathfrak{A}, T) &= \mathbb{E}_{\sigma \in \mathcal{P}_l}\left[\text{REG}_{I_\sigma}(\mathfrak{A}, T)\right] \\ &= \frac{1}{|\mathcal{P}_l|} \cdot \left(\sum_{\sigma \in \mathcal{P}_l} \text{REG}_{I_\sigma}(\mathfrak{A}, T)\right). \end{aligned} \tag{5.1}$$

Given the random permutation regret, we define the following property of algorithm, which will play an essential role in obtaining truthful algorithm in the single-agent setting.

Definition 5.4.2 (Truthful under random permutation)**.** Given an arbitrary single-agent bandit instance $\mathcal{I}$ with l arms, consider a truthful strategy $\mathcal{O}$ and arbitrary strategy $\mathcal{S}$. An algorithm $\mathfrak{A}$ is *truthful under random permutation* (TRP) if

$$\text{RP-REG}(\mathfrak{A}, T; \mathcal{O}) \leq \text{RP-REG}(\mathfrak{A}, T; \mathcal{S}).$$

Note that TRP requires the inequality to holds for arbitrary bandit instance $\mathcal{I}$. The foundation of these notions essentially builds upon the observation from our counterexample provided in Theorem 5.3.1. To elaborate more, consider a strategy $\mathcal{S}$ that only replicates the first arm given a bandit instance $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_l$. Then, we can observe that any permutation $\sigma \in \mathcal{P}_l$ results in a bandit instance that belongs to the following family of bandit instances: define $\mathcal{I}^* = \cup_{i \in [l]} \mathcal{I}_i$ such that

$$\mathcal{I}_1 = \{\underbrace{\mu_1, \mu_1}_{repl}, \mu_2, \ldots, \mu_l\}, \ldots, \mathcal{I}_l = \{\mu_1, \mu_2, \ldots, \underbrace{\mu_l, \mu_l}_{repl}\}.$$

In words, $\mathcal{I}_i$ denotes the case that the strategy $\mathcal{S}$ replicates the arm with i-th highest reward. For example, suppose that $\sigma \in \mathcal{P}_l$ satisfies $\sigma(1) = i$ for some $i \in [l]$. Since $\mathcal{S}$ replicates the first arm,

under the permutation σ, the arm with parameter μ_i is replicated, and thus it corresponds to $\mathcal{I}_i$. Then, our notion of TRP essentialy asks what is an algorithm that makes the expected regret of $\mathcal{I}$ smaller than (or equal to) the average of expected regret of $\mathcal{I}_1, \ldots, \mathcal{I}_l$. Put it differently, TRP essentially asks the following fundamental question, as we pointed out in the introduction.

Given a bandit instance, does "uniform randomly" replicate one arm increase the expected regret?

This foundation of TRP will play an essential role in constructing replication-proof algorithm in both the single-agent and the multi-agent setting. Indeed, from our proof of Theorem 5.3.1, one can observe that UCB does not satisfy the above question, and thus fails to be replication-proof.

Definition 5.4.3 (Permutation invariance)**.** An algorithm is *permutation invariant* (PI) if given the random bits of the algorithm and the arms, the choice of which arms to pull remains the same (up to tie-breaking) when the index of the arms are permuted.

Note that the permutation-invariance essentially requires the algorithm to be agnostic to the index of the arms, *i.e.,* does not use the arm index to choose the arm, possibly except the tie-breaking case. One can easily verify that it holds for a broad class canonical algorithms such as UCB and ε-greedy.

Theorem 5.4.4. *In the single-agent setting, if an algorithm is TRP and PI, then it is replication-proof.*

The proof essentially follows from the observation that the algorithm's ex-ante regret can be partitioned into a disjoint set of random permutation regret, with respect to the ex-post realization of the mean rewards. Based on this disjoint partition, we can effectively marginalize over the priors while maintaining the partition, and conclude that the dominance over the partition yields the overall dominance.

Furthermore, we observe that TRP is indeed a necessary condition for an algorithm to be replication-proof in the single-agent setting.

Proposition 5.4.5. *If an algorithm is not TRP, then it is not replication-proof.*

Discussion 5.4.6. One may wonder if PI can be violated for any replication-proof algorithm. For example, one can consider a black-box algorithm that first randomly selects an agent without using any statistical information but only with the agent indices, and runs a bandit algorithm within the agent. Further, suppose that it asymmetrically favors some agents' indices with higher probability in the first phase. Such an algorithm indeed is not permutation invariant, since the permutation of agent index would lead to a different outcome given the same reward tapes. From agents' perspective, however, each agent only needs to care about the secondary arm selection phase after the agent selection since they cannot change the probability to be selected in the agent selection phase. Using these observations, one can observe that such an algorithm is indeed replication-proof, but not permutation invariant. One may refine the notion of PI to be restricted within each agent, but we do not argue more details as it is beyond of our interest.

Constructing TRP algorithm. We now present an algorithm that satisfies TRP. In what follows, we mainly prove that exploration-then-commit (ETC) satisfies TRP once equipped with a proper parameter. Its pseudocode is presented in Algorithm 3 in the appendix. We suppose that the algorithm breaks tie in a uniform random manner.[8] ETC essentially decouples the exploration phase and the exploitation phase very explicitly, and thus brings a significant advantage in comparing its expected regret under several problem instances.

For analytical tractability, we pose some *assumptions* on the support of the prior distributions. Namely, for any agent i, F_i has a discrete support over $[0, 1]$ for $i \in \mathcal{N}$. Further, we define Δ_i be the minimum gap between any two possible outcomes from F_i, and let $\Delta = \min_{i \in \mathcal{N}} \Delta$. Furthermore, the algorithm knows this gap Δ.

[8] Any tie-breaking rule does not hurt our analysis.

Discussion 5.4.7. We discuss how one can weaken these assumptions at the end of this section, possibly at the cost of polylogarithmic blowup in the regret rate or another assumptions. We also remark that such assumptions, especially that the algorithm knows the minimum gap, often appear in the literature, *c.f.,* [32, 117, 118].

Discussion 5.4.8. In practice, such scenario is fairly plausible in real-world applications. For example in content platforms, suppose that a content creator comes up with several contents and is about to register. Assume the reward of each content is just the number of views. Typically, the content creator does not exactly know the reward of each content but has historical data on its previous contents. The historical data can be used to structure a prior distribution of the quality of his/her contents. Since each content creator usually has a very small number of original contents compared to the recommendation algorithm's time scale (one for each traffic), this prior distribution may consist of a fairly limited number of samples, thereby inducing a discrete support with a few number of points.

Then, our main result in the single-agent setting can be written as follows.

Theorem 5.4.9. *ETC with exploration length* $m \geq \frac{2}{\Delta^2} l \ln(2T)$ *is TRP in the single-agent setting. Further, it has regret upper bound of*

$$\sum_{a \in [l]} \Big(\frac{2\delta_a l \ln(2T)}{\Delta^2} + 1 \Big),$$

where l *denotes the number of arms and* δ_a *denotes the gap of the mean rewards between optimal arm and arm* a *for the single agent.*

The proof essentially follows from the standard regret analysis of ETC along with carefully chosen m. By setting m sufficiently large, we can essentially decrease the probability that a suboptimal

arm is chosen to be $o(1/T)$. This will guarantee that the expected number of rounds the suboptimal arm is chosen is smaller than 1, whereas any replicating strategy incurs a simple regret lower bound larger than this quantity. Together with Theorem 5.4.4 and the fact that ETC is PI (which is straightforward to verify), it follows that ETC is replication-proof in the single-agent setting.

Note that the proof of Theorem 5.4.9 heavily relies on the fact that the algorithm can observe Δ in advance and set the exploration length correspondingly. Thus, despite we aim to obtain a *prior-independent* algorithm that does not require any knowledge in the agent's prior distributions, it is not *truly* prior-independent in an algorithmic manner since it requires some problem parameters to operate the algorithm. We discuss several ways to refine such restrictions, thereby suggesting a road for truly prior-independent algorithm. First, assume Δ is constant but not known to the algorithm. In this case, one may replace Δ in the algorithm to be some increasing functions $f(T) = \omega(1)$, and it is straightforward to verify that Theorem 5.4.9 still holds, but in an asymptotical regime. Similarly, one can effectively wipe out the dependency on the total number of original arms l by assuming that $l = O(1)$ and replacing it with some increasing functions. Furthermore, the knowledge on T can be wiped out by the standard doubling trick, *c.f.,* see Chapter 1 in [103], without sacrificing the truthfulness. Remark that all these refinements only introduce polylogarithmic blowup in the regret upper bound, where we do not provide a formal proof as it is a cumbersome application of standard techniques. The following theorem spells out these arguments.

Proposition 5.4.10. *Assume $l = O(1), \Delta = \Omega(1)$. Then, there is a prior-independent algorithm with polylogarithmic regret satisfying TRP and PI in the single-agent setting for sufficiently large T, that does not require any information on problem parameters.*

5.5 Multi-agent Replication-proof Algorithm

Now we turn our attention to the more general setting with multiple agents. We start by using the machinery of *hierarchical structure* in H-UCB, but in a different manner. A direct implication of Theorem 5.4.4 is that if we run a variant of H-UCB such that the first phase runs a simple uniform-random selection algorithm, and the second phase runs some arbitrary TRP and PI algorithms, then the resulting algorithm is replication-proof. To formally see this why, let $\mathcal{S}_i$ be the strategy of agent i. Define $\Gamma_i(\mathcal{S}_i)$ be the agent i's utility under the single-agent setting with agent i, *i.e.,* when there are no other agents other than i. Due to the uniformly random selection nature in the first phase, the expected utility of any agent i can be simply written $u_i(\mathcal{S}_i, \mathcal{S}_{-i}) = \Gamma_i(\mathcal{S}_i)/n$. Note that the dynamic of the algorithm purely reduces to the dynamic of the second phase algorithm if there's only a single agent. Hence, by Theorem 5.4.4 and due to the TRP and PI of the second phase algorithm, we conclude that truthful registration is a dominant strategy for any agent i, and thus the result follows. Note, however, that if we deploy uniform random selection algorithm in $\textsc{Alg}_1$, it essentially suffers *linear regret* since it always choose any suboptimal agent with constant probability at every rounds.

To capture both of sublinear regret and replication-proofness, we need more sophisticated algorithm with sublinear regret in the first phase. To this end, we present H-ETC-R presented in Algorithm 5 in the appendix, which adopts ETC in both phases along with additional device of *restarting round*. Similar to H-UCB, this consists of two phases each running ETC agent-wise and arm-wise manner, respectively. For analytical tractability, we say that the set of arms belong to *stochastically ordered family* if two arms a and b satisfy $\mu_a \geq \mu_b$, then the reward distribution of arm a is (first-order) stochastic dominant over that of b, *i.e.,* $\Pr[r_a \geq x] \geq \Pr[r_b \geq x]$ for any $x \geq 0$, where r_i denotes arm i's

reward random variable.[9] Further, define $L = \max_{i \in [n]} l_i$. Then, our main result is as follows.

Theorem 5.5.1. *Consider a stochastically ordered family of arms, and discrete support of prior distributions. Consider* H-ETC-R *with* $M \geq mL$ *and* $m = \frac{2}{\Delta^2} L \ln(2T)$*, and* $\tau = Mn$ *Then,* H-ETC-R *is replication-proof.*[10]

Proof sketch. We provide a brief proof sketch. The proof mainly relies on the single-agent result in Theorem 5.4.4 along with a number of algebraic manipulations which heavily relies on the nature of ETC and the restarting round. Let us focus on agent i. We can decompose agent i's total expected utility to be (i) exploitation phase utility and (ii) exploration phase utility. Given a realization of the reward tapes, the utility from (i) and (ii) is independent due to the nature of ETC. Since (i) simply reduces to running single-agent ETC for each agent, we can compare (i) of two strategies using the single-agent result, thus truthful strategy is at least better for (i) thanks to the single-agent theorem.

For (ii), if agent i is not selected in the exploitation phase, it is simply zero, so we can only focus on the case when agent i is selected in the exploitation phase. Thus, its expected utility from (ii) can be written as multiplication of (a) probability that agent i has the largest empirical average reward and (b) the expected utility by being pulled over the exploitation rounds. For (b), since our choice of τ *restarts* Phase 2 for any agent exactly at the beginning of exploitation phase of Phase 1, the expected utility therein simply reduces to the expected utility of single-agent setting with agent i given the time horizon $T - Mn$. Thus, again by our single-agent result of Theorem 5.4.4, the truthful strategy yields larger (or equal) quantity for (b) as well. Finally, by carefully using a series of concentration inequalities, we bound the difference from (a) between the truthful strategy and any other strategy is relatively smaller

[9] For example, Bernoulli arms and Gaussian arms parameterized by mean rewards, belongs to this family. We also refer to [119] for more examples in exponential family.

[10] Similar to the discussion in Proposition 5.4.10, we can refine this results and make the algorithm truly prior-independent. We omit these details of beyond the interest.

and by doing so, loss from (a) cannot make up the gains from (b) and (i), which eventually implies the dominance of truthful strategy. a

We finalize our results by presenting regret upper bound of the H-ETC-R with ETC and ETC, which concludes that it achieves sublinear regret as well as replication-proofness.

Theorem 5.5.2. *Set* ALG_1 *be ETC with* $M = \max(mL, \sqrt{T \ln T})$*, and* ALG_2 *be ETC with* $m = {}^2/_{\Delta^2} \cdot L \ln(2T)$*. Then,* H-ETC-R *has expected regret of* $O(\frac{nL^3\sqrt{T \ln T}}{\Delta^3})$.

The regret analysis is based on a construction of clean event on each agent under which the agent's empirical average reward becomes close enough to his own optimal arm's mean reward. The deviation probability of this clean event can be obtained using tail bounds along with the regret analysis of ETC, which essentially implies the necessity of having ALG_1's exploration length polynomial. The regret bound then easily follows from obtaining the probability that all the clean event holds. The formal proof is deferred to appendix.[11]

[11] We remark that the algorithm's dependency on the problem parameters can be weakened similar to Proposition 5.4.10, but we omit the details as it goes beyond our interest.

Part III

Delegation of Trade

In this part of the article, I will present our results on the fundamental problem called *bilateral trade*. Each chapter presents a self-contained paper, which is either finalized or a work in progress.

The problem of bilateral trade, introduced by [14], has been a cornerstone in mechanism design and algorithmic game theory in the past few decades. In bilateral trade, a single buyer and seller engage in trade, with each party holding private information about their valuation. If the buyer's valuation is larger than the seller's, social welfare strictly increases by transferring the item from the seller to buyer at certain prices. The buyer and seller, however, may strategically hide their true valuation to increase their own payoff. This leads to the challenge of constructing a truthful mechanism while maintaining its efficiency in terms of the social welfare. More precisely, the objective here is to design a mechanism that optimizes efficiency while satisfying (i) incentive-compatibility (IC), *i.e.,* both players truthfully report their values, (ii) individual rationality, *i.e.,* the players are not worse off by participating the trade, and (iii) budget-balancedness (BB) so that the mechanism does not run a deficit.

The seminal result by [14], however, reveals that it is impossible to achieve ex-post efficiency[12] subject to the constraints of IC, IR, and BB, and even subject to the loosened notion of Bayes-Nash IC (BNIC), IR, and weakly budget-balancedness (WBB). Correspondingly, a long line of work [120–123] has studied the best possible approximate efficiency, in particular for the notion of *gains-from-trade* (GFT),[13] with respect to those desideratum. It was a long-standing open problem whether a constant-factor approximation to the first-best GFT was possible, until it was recently answered in the

[12] A mechanism is ex-post efficient if it achieves the optimal ex-post social welfare, *i.e.,* if trade occurs whenever the buyer's valuation is at least the seller's valuation.

[13] Gains-from-trade is defined by the expected marginal increase of the social welfare. Thus, it is typically harder to approximate than social welfare. We refer to Section 6.2 for more details.

affirmative by [15].

Chapter 6: Gains-from-trade in Bilatreral Trade with a Broker

This chapter presents our results on bilateral trade problem with a broker [17].

6.1 Introduction

In many real-world scenarios, the buyer and seller might not be able to trade the item directly, *e.g.,* due to regulatory requirements, privacy concerns, or time constraints, in particular for financial markets, real estate, over-the-counter markets, and markets for demerit goods with licensed intermediaries. Brokers do more than just facilitate transactions, where they often play a crucial role in price discovery and risk management. Notably, [14] also study bilateral trade with a broker and characterizes an optimal mechanism for a broker who tries to maximize her own profit. Despite being significantly practical and well-founded, there have only been a few papers studying bilateral trade in the presence of broker [124–127].

In this work, we initiate the study of approximating gains-from-trade in bilateral trade with a broker, investigating the following fundamental question, which has not been explored before:

Does the presence of a strategic broker interfere with the possibility of constant-factor approximation to the first-best gains-from-trade?

We answer this question both affirmatively and negatively by characterizing the regimes under

which constant approximation is possible or not. Our characterization reveals the extent to which the marginal increase of the social welfare can be interrupted by a strategic broker, shedding light on the public decision-making process of enforcing regulatory restriction in trade.

6.1.1 Our Results

To summarize our results, we first briefly introduce the problem setup. In bilateral trade with a broker, there exist a buyer and a seller with private valuations v and c, respectively. We use F and G to denote the respective cumulative probability distributions (CDFs) of the buyer and the seller, and f and g to denote their respective probability density functions (PDFs). In addition, we write $\mu_F(\alpha)$ for $\alpha \in [0,1]$ to denote the α-quantile of F, *i.e.,* $\mu_F(\alpha) = \inf\{t : F(t) \geq \alpha\}$, and similarly define $\mu_G(\beta)$ for $\beta \in [0,1]$ to be the β-quantile of G. The buyer and the seller exclusively trade through a broker, who tries to maximize her expected profit by constituting an arbitrage on the prices to the buyer and seller. We consider *Bayesian mechanisms* in which the broker knows F and G.[1] The gains-from-trade in this setting is defined to be the expected marginal increase of the social welfare, *i.e.,* by $\text{GFT} = \mathbb{E}_{v\sim F, c\sim G}\left[(v-c)\cdot x(v,c)\right]$, where $x(v,c)$ denotes the probability that the item is transferred from the seller to the buyer if their respective valuations are c and v. If there exists an omniscient planner who can observe the private valuations and also enforce trade at arbitrary prices, society would enjoy the maximum increase of social welfare, *i.e.,* the first-best GFT (henceforth FB-GFT), defined as $\text{FB-GFT} = \mathbb{E}_{v\sim F, c\sim G}\left[(v-c)\cdot \mathbf{1}\left\{v \geq c\right\}\right]$. We refer to Section 6.2 for more detailed definitions and notation.

Distribution Types. We say that the buyer's distribution is *regular* if his virtual valuation $\phi_F(x) =$

[1] In comparison, the literature often considers *prior-independent* mechanisms that do not know the distributions, *e.g.,* second-price auctions.

$x - (1 - F(x))/f(x)$ is nondecreasing, and we say that the seller's distribution is regular if his virtual valuation $\phi_G(x) = x + G(x)/g(x)$ is nondecreasing. The quantities $h_b(x) = f(x)/(1 - F(x))$ and $h_s(x) = g(x)/G(x)$ are usually denoted as *hazard rates*. We say that F has a *monotone hazard rate* (MHR) if $h_b(x)$ is nondecreasing, and we say that G is MHR if $h_s(x)$ is nonincreasing. MHR and regular distributions are widely studied in the literature, with their fruitful connections to mechanism design ([122, 123, 128]). Finally, we often consider the ex-ante *symmetric* agents setting, in which the buyer's and seller's distributions are the same, *i.e.,* $F = G$.[2]

Posted-Pricing Mechanisms. We begin by investigating a broker who runs a *posted-pricing* mechanism,[3] *i.e.,* a broker who offers a pair of fixed prices p and q to the buyer and seller, respectively. Posted-pricing mechanisms are widely adopted in many real world applications due to their simplicity and transparency, in particular for online labor market and size-discovery mechanisms, including workup mechanisms and dark pools ([129]). In addition, in the standard setting of bilateral trade without a broker, posted-pricing mechanisms were shown by [133] to be the only class of mechanisms satisfying budget-balancedness and DSIC. Note further that posted-pricing mechanisms are clearly IR.

Towards obtaining approximation factors to FB for various regimes of problem instances, we first identify a structural inequality involving GFT, the broker's expected profit, and FB-GFT, which might be of independent interest. To this end, we use a variant of the analysis by [120] after decomposing the gains-from-trade using the posted prices.

Theorem 6.1.1 (Decomposition)**.** *Given a broker's posted-pricing mechanism with prices* (p, q)*, the*

[2]Such ex-ante symmetric agents are well-motivated by size-discovery mechanisms in financial markets – in particular for dark pools [129] – and are widely assumed in the literature, *e.g.,* see [130–132].

[3]This is often called a fixed-price mechanism in the literature.

following holds:

$$\text{GFT} \geq \textsc{Profit} + \min(G(q), 1 - F(p)) \left(\text{FB-GFT} - \int_q^p G(x)(1 - F(x))\, dx \right). \tag{6.1}$$

As a byproduct, we reobtain the structural result by [120], including the $1/2$-approximation to the first-best GFT via median-based pricing. We state this formally in the following corollary:

Corollary 6.1.2. *Consider bilateral trade without a broker*[4] *and a posted-pricing mechanism with fixed price p. Then, we have*

$$\text{GFT} \geq \min(G(p), 1 - F(p)) \cdot \text{FB-GFT}.$$

As a special case, if the median of F is at least that of G, then setting p to be any value between these medians yield a $1/2$-approximation to the first-best GFT.

Note that Corollary 6.1.2 immediately follows from plugging $p = q$ into Theorem 6.1.1, which results in $\textsc{Profit} = 0$ and $\int_q^p G(x)(1 - F(x))\, dx = 0$,

On the other hand, it is not immediately clear what approximation factors we can obtain from Theorem 6.1.1 when a profit-maximizing broker is in fact present. Let (p^*, q^*) be the broker's optimal pair of prices. As an initial attempt, suppose that the broker uses the $1/4$-quantile of the seller's distribution and the $3/4$-quantile of the buyer's distribution as prices for the seller and buyer, respectively, assuming that $\mu_F(3/4) \geq \mu_G(1/4)$. For simplicity, let $p = \mu_F(3/4)$, $q = \mu_G(1/4)$, $\Delta = p - q$ and $H(p, q) = \int_q^p G(x)(1 - F(x))\, dx$. Then, to lower bound the right-hand side of Theorem 6.1.1, we might naively upper bound the integral by $\Delta \cdot \max(G(p), 1 - F(q))$. However, this involves the gap

[4] Alternatively, a broker who tries to maximize the gains-from-trade instead of her expected profit.

Δ between the posted prices as well as the probabilities $G(p)$ and $1 - F(q)$, none of which are easy to deal with.

Instead, taking a closer look at (6.1), we consider two cases where $H(p, q) = \int_q^p G(x)(1-F(x))$ is sufficiently small or moderately large, respectively. If $H(p, q)$ maintains a rather small value, then since

$$\textsc{Profit} = (p - q)(1 - F(p))G(q) \leq \int_q^p (1 - F(x))G(x)\, dx,$$

it should be the case that PROFIT is also small. Thus, there is a high chance that $\min(G(q), 1 - F(p))$ is small as well, *i.e.,* we cannot find an appropriate lower bound of the right-hand side of (6.1) in multiplicative terms of FB-GFT.

On the other hand, suppose $H(p, q)$ is moderately large, say $H(p, q) \geq C \cdot$ FB-GFT for some constant $C > 0$. Then, since the profit-maximizing broker's optimal payoff should be at least the payoff obtained from posting (p, q), we know that

$$\textsc{Profit}(p^*, q^*) \geq \textsc{Profit}(p, q) \geq \Delta \cdot \frac{1}{16},$$

where we obtain the $1/16$ due to our choice of (p, q) above. At the same time, Δ provides a trivial upper bound on $H(p, q)$, which implies that we can lower bound the optimal broker's profit by $C' \cdot$ FB-GFT for some constant $C' > 0$. The main technical challenge here is to prove that there exists an appropriate choice of a pair of prices (p, q) such that the corresponding quantile probabilities are absolute constants and $H(p, q)$ is at least some constant times FB-GFT. We show that this is indeed possible for MHR distributions with symmetric agents, using a clever decomposition and an exploitation of the properties

of MHR distributions. Formally, we prove the following theorem:

Theorem 6.1.3 (Symmetric agents, MHR distributions)**.** *Consider symmetric agents with MHR distributions. Then, if the broker runs an optimal posted-pricing mechanism,*[5] *it follows that* GFT $\geq$ $1/39 \cdot$ FB-GFT.[6]

It is worth noting that symmetric agents are well-motivated by size-discovery mechanisms including workup mechanisms by the U.S. Treasury market and dark pools [129, 131, 132]. In particular, in stock markets with brokerage, each participant can act either as a buyer or a seller. Thus, the broker cannot distinguish between the buyer and seller and hence cannot estimate their respective distributions separately. MHR distributions are also central to the theory of mechanism design along with regular distributions, due to the celebrated result by [128] and further seminal works in algorithmic game theory [34, 134–136] (see [137] for a comprehensive exposition).

To tackle the relatively small approximation guarantee presented above, we show that if the hazard rates of the MHR distributions are further bounded above by $M > 0$, then any optimal posted-pricing mechanism achieves at least a $2/M^2$-approximation. The proof follows from expanding FB-GFT and GFT in a slightly different manner than in Theorem 6.1.1, largely by exploiting symmetry and properties of MHR distributions. The formal statement can be presented as follows:

Theorem 6.1.4 (Symmetric agents, MHR distributions with bounded hazard rates)**.** *Consider symmetric agents with MHR distributions whose hazard rates are uniformly bounded above by M. Then, if the broker runs an optimal posted pricing mechanism, it follows that* GFT $\geq 2/M^2 \cdot$ FB-GFT.[7]

[5] Note that the result here only holds for an optimal posted-pricing mechanism, not an arbitrary posted-pricing mechanism. Indeed, if the broker is adversarial rather than profit-maximizing, one cannot obtain any approximation factor since she may give up her profit but try to ruin society by making no trade at all.

[6] Note that Theorem 6.1.3 immediately implies that under symmetric agents with MHR distributions, any optimal posted-pricing mechanism achieves a $1/39$-approximation to the first-best social welfare.

[7] Note that M is guaranteed to be at least 2 in such problem instances, implying that the best possible approximation factor here is $1/2$.

So far, we have only considered optimal posted-pricing mechanisms, which requires the broker to have complete information, *i.e.,* a precise estimation, about the distributions F and G. In practice, however, it is typically difficult for the broker to exactly estimate the probability distributions of the buyer and the seller. Instead, the broker may only be aware of several statistics about the distributions such as expectations, medians, or more generally quantiles. Therefore, our next question is to characterize the approximation factor if the broker runs a *quantile-based* posted-pricing mechanism where she offers $\mu_F(\alpha)$ to the buyer and $\mu_G(\beta)$ to the seller. In this case, we observe that one can express the approximation factor to the first-best GFT as a function of α and β without introducing the gap between the quantiles itself, formalized as follows:

Theorem 6.1.5 (Quantile-based posted-pricing mechanism)**.** *Suppose* $\mu_G(\beta) \leq \mu_F(\alpha)$, *with* $\alpha < 1$ *and* $\beta > 0$. *Then, if the broker sets* $p = \mu_F(\alpha)$ *and* $q = \mu_G(\beta)$, *we have that*

$$\text{GFT} \geq \beta(1-\alpha)\text{FB-GFT}.$$

Alternatively speaking, since we know that $\textsc{Profit} = (\mu_F(\alpha) - \mu_G(\beta))(1-\alpha)\beta$, this is equivalent to $\text{GFT} \geq \textsc{Profit}/\Delta \cdot \text{FB-GFT}$, where Δ denotes the gap between the prices. As an example of Theorem 6.1.5, if the broker sets the $3/4$-quantile (*i.e.,* the third quartile) of the buyer's distribution as the price for the buyer and the $1/4$-quantile (*i.e.,* the first quartile) of the seller's distribution as the price for the seller, then the gains-from-trade is at least $1/16$ of the first-best GFT.

It is also worth noting that Theorem 6.1.5 immediately implies a weaker version of the result by [120]. Indeed, assuming that $\mu_G(1/2) \leq \mu_F(1/2)$, plugging in $\alpha = \beta = 1/2$ yields a $1/4$-approximation to the FB-GFT, which is strictly weaker than Corollary 6.1.2. Conceptually, therefore, Theorem 6.1.5 completely characterizes the approximation factor using the probabilities that corre-

spond to the quantiles while sacrificing a reasonable amount of loss in the approximation factor.

An interesting implication of this result is that under symmetric agents, a posted-pricing mechanism that draws a single sample from each of the (identical) distributions of the buyer and the seller and offers the larger sample as the price to the buyer and smaller sample as the price to the seller achieves a $1/12$-approximation to FB-GFT. In other words, if the broker has limited information, then society can enjoy a constant fraction of the first-best gains-from-trade, even without any distributional assumptions such as MHR.

Corollary 6.1.6. *Suppose $F = G$. Then, given $p \sim F$ and $q \sim G$, the posted-pricing mechanism $(\max(p, q), \min(p, q))$ obtains a $1/12$-approximation to the first-best gains-from-trade in expectation.*

(In)approximability beyond Posted-Pricing Mechanisms. So far, we have restricted the broker to running a posted-pricing mechanism. In the real world, however, the broker may run any arbitrary mechanism to maximize her profit. Thus, we now consider the case where the broker runs any BNIC and IR mechanism that maximizes her profit. The seminal result by [14] characterizes one such optimal mechanism, which even turns out to be DSIC:

Theorem 6.1.7 ([14])**.** *Suppose the buyer's distribution F and the seller's distribution G are regular with bounded supports. Then, among all BNIC and IR mechanisms, the broker's expected profit is maximized by a mechanism that transfers the item from the seller to the buyer if and only if $\phi_F(v) \geq \phi_G(c)$.*

Note that due to the strategic behavior of the broker, the above mechanism may result in strictly less trading than a mechanism that maximizes GFT. Indeed, if one considers the symmetric setting where both buyer's and seller's valuations have uniform distribution $U[0, 1]$, one can observe that under the above mechanism, trade occurs if and only if $\phi_F(v) = 2v - 1 \geq \phi(c) = 2c$, *i.e.,* if and

only if $v - c \geq 1/2$. However, as shown by [128], a BNIC and IR mechanism that maximizes the gains-from-trade in this example transfers the item if and only if $v - c \geq 1/4$. Thus, the existence of the broker interferes with social welfare. For uniform distributions, however, we show that a $1/2$-approximation to the first-best GFT is still possible, and that this bound is even tight.

Theorem 6.1.8 (Uniform distribution)**.** *If F and G are uniform distributions, then* GFT $\geq$ FB-GFT$/2$ *for any BNIC and IR mechanism that maximizes the broker's expected profit. Furthermore, this bound is tight in the sense that there exists a problem instance in which the approximation ratio is exactly* $1/2$.

The main technical subtlety in proving Theorem 6.1.8 lies in handling a number of cases depending the supports of the uniform distributions. To simplify the possible cases here, we introduce a result stating that GFT$/$FB-GFT is invariant to stretches and shifts in the buyer's and seller's distributions as long as we apply the same transformations to both distributions, formally presented in Lemma 6.4.1.

On the other hand, if we consider general distributions, the extent to which the broker's strategic behavior degrades the GFT can be significant, formally stated as follows:

Theorem 6.1.9 (Inapproximability)**.** *For any $\varepsilon > 0$, there exists a problem instance such that* GFT $<$ $\varepsilon \cdot$ FB-GFT *for any BNIC and IR mechanism that maximizes the broker's expected profit.*

The proof here is based on the characterization in Theorem 6.1.7. To briefly summarize, we start by trying to come up with a regular distribution for the buyer such that $\phi_F(v) < \phi_G(c)$ for all v and c, *i.e.,* such that trade never occurs under any BNIC and IR mechanism that maximizes the broker's expected profit. Solving a resulting differential equation, we obtain $F(x) = -a/x + 1$ for some constant $a > 0$. This distribution, however, has an unbounded support of $[a, \infty)$, so we cannot guarantee the optimality of the mechanism given in Theorem 6.1.7, which only applies to

bounded supports. Indeed, one can see that the above choice of F results in PROFIT $= 0$, whereas a simple posted-pricing mechanism can achieve a strictly positive expected profit (assuming the problem instance is nontrivial, *i.e.,* FB-GFT > 0). Thus, we need to truncate this distribution carefully to obtain the desired result.

One natural follow-up question to Theorem 6.1.9 is to what extent such inapproximability persists. We first consider a restricted set of problem instances where exactly one of the buyer and seller has public value. Formally speaking, *public-buyer* refers to the case where the buyer's ex-post valuation v is deterministically known to the broker, and similarly for *public-seller*. We require analogues of Theorems 3 and 4 in [14] to characterize an optimal BNIC and IR mechanism tailored to these specific problem instances, since we cannot directly apply [14] when there are probability distributions with singleton support. These lemmas yield the following inapproximability result:

Theorem 6.1.10 (Inapproximability with public seller or public buyer)**.** *For any $\varepsilon > 0$, there exists (i) a continuous distribution F and a singleton distribution G and (ii) a singleton distribution F and a continuous distribution G such that* GFT $< \varepsilon \cdot$ FB-GFT *for any BNIC and IR mechanism that maximizes the broker's expected profit.*

Furthermore, we observe that inapproximability also carries over when the agents are symmetric.

Theorem 6.1.11. *For any $\varepsilon > 0$, there exist identical distributions $F = G$ such that* GFT $< \varepsilon \cdot$ FB-GFT *for any BNIC and IR mechanism that maximizes the broker's expected profit.*

Finally, we show that in the general-distribution and public-seller settings, the inapproximability of the first-best gains-from-trade carries over to the first-best social welfare as well.

Theorem 6.1.12. *For any $\epsilon > 0$, there exists distributions F and G such that* SW $< \epsilon$FB-SW *for*

any BNIC and IR mechanism that maximizes the broker's expected profit. Furthermore, this statement holds even if we restrict G to be deterministic.

6.1.2 Further Related Work

Bilateral Trade. The theory of bilateral trade dates back to the seminal result by [14], where the authors proved that it is impossible to have an ex-post efficient mechanism that satisfies BNIC, IR and WBB. Since then, there has been a number of attempts to identify the best possible approximation ratio to the first-best gains-from-trade. [122] and [123] prove that a constant approximation ratio is possible for certain classes of distributions, *e.g.,* regular and MHR, and provide some worst-case upper bounds on the approximation ratios. More recently, [15] provide a mechanism that achieves a $1/8.23$-approximation to the first-best gains-from-trade. [138] significantly improves this to $1/3.15$ and further characterizes the exact approximation factor of the seller-pricing mechanism, given that the buyer's distribution satisfies MHR. Most recently, [139] study the problem of approximating gains-from-trade and social-welfare for bilateral trade when the buyer's and seller's valuations can be correlated. Finally, another line of literature in bilateral trade focuses on identifying the best-possible approximation ratio for the class of posted-pricing mechanisms [129, 140, 141]. It is worth noting that some of the above work focuses on approximating the social welfare defined by $\mathbb{E}_{v\sim F, c\sim G}\left[c + (v-c)\cdot \mathbf{1}\left\{ v \geq c \right\}\right]$, which is much easier to approximate than gains-from-trade, since social welfare attains positive value even if trade does not occur.

Bilateral Trade with a Broker. In comparison to the significant developments in bilateral trade, bilateral trade with a broker has been less studied, despite its significance and even its introduction in the same seminal paper by [14]. In terms of prior work on this problem, [14] characterize one

BNIC and IR mechanism that maximizes the broker's expected profit. [125] studies the impact of an imperfectly-informed broker who has partial information about the buyer's and seller's values, *e.g.,* the broker knows the the bounded interval in which the values are contained. [127] consider a variant of the standard bilateral trade setting in which a mediator commits to a mechanism that possibly includes multiple rounds of private communication and runs a posted-pricing mechanism after the communication ends. Notably, their setting does not consider prior information on the buyer's and seller's distributions, *i.e.,* their setting considers prior-independent mechanisms. Thus, certain information should be exchanged to induce an efficient trade, and the solution concept therein is a perfect Bayes equilibrium. In a similar vein, [124] study robust prior-independent mechanisms when the buyer's and seller's valuations are discounted over time and hence both agents are eager to have the trade occur as soon as possible. In this setting, the mediator keeps the reported valuations of the buyer and the seller privately while trade is incompatible. Then, after trade becomes compatible, the mediator discloses the agreement and trade occurs at the agreed price. We emphasize that none of the literature has studied the approximation ratio for social welfare or gains-from-trade in the standard bilateral trade with broker model introduced by [14].

Two-sided Markets. Several works consider an extension of bilateral trade that is typically referred to as two-sided markets or double auctions. In this setting, there are multiple sellers and buyers, each of which is interested in selling his own items and buying some items of interest. The seminal result by [142] presents a double-auction, trade-reduction mechanism that is IR, DSIC, BB, and achieves a high ex-post efficiency under a large market. Since then, several works have studied prior-independent mechanisms for double-auction settings [143, 144]. [19] and [145] study the Bulow-Klemperer [16] result for gains-from-trade in double-auction settings, analyzing the power of recruiting more buyers

and sellers for prior-independent mechanisms, and [146] and [147] study the approximate mechanism for bilateral trade in double auctions in both the single-dimensional and multi-dimensional settings.

6.2 Problem Setup

Bilateral Trade. In the bilateral trade problem, a seller tries to sell an item to a buyer. Both the buyer's and the seller's valuations are private and independently drawn from cumulative distributions (CDFs) F and G, respectively, both of which are supported on $\mathbb{R}_{\geq 0}$. Following [14], we write v and c to denote the buyer's and the seller's respective valuations. We write f and g to denote the density functions of F and G, respectively.

If trade occurs between the buyer and the seller with valuations v and c, respectively, then society realizes an increment of welfare of amount $v - c$. If the allocation function of a mechanism (probability of trade) is $x(\tilde{v}, \tilde{c})$ given the reported valuations $\tilde{v}$ of the buyer and $\tilde{c}$ of the seller, then the *gains-from-trade*[8] (GFT) is defined by the expected amount of increased utility from trade, *i.e.,*

$$\text{GFT} = \mathbb{E}_{v \sim F, c \sim G} \left[(v - c) \cdot x(v, c) \right],$$

assuming truthful reporting of the buyer and the seller. If there exists an omniscient social planner who can observe the private valuations, then he can enforce trade whenever the buyer's valuation is at least the seller's valuation. Such an ideal mechanism induces $x(v, c) = \mathbf{1}\{ v \geq c \}$, and the resulting GFT is defined as the *first-best gains-from-trade* (FB-GFT), *i.e.,*

$$\text{FB-GFT} = \mathbb{E}_{v \sim F, c \sim G} \left[(v - c) \mathbf{1}\{ v \geq c \} \right].$$

[8] Note that this differs from social welfare since social welfare is c even if trade does not occur.

Similarly, the *social welfare* (SW) is defined by the expected utility from trade, *i.e.,*

$$\text{SW} = \mathbb{E}_{c\sim G}[c] + \text{GFT},$$

and the *first-best social welfare* (FB-SW) is defined as the social welfare of the ideal mechanism described above, *i.e.,*

$$\text{FB-SW} = \mathbb{E}_{c\sim G}[c] + \text{FB-GFT}.$$

A mechanism is *ex-post efficient* if it achieves the first-best GFT. However, the seminal result by [14] asserts that it is not possible to construct a mechanism that is simultaneously (i) Bayesian-Nash incentive-compatible (BNIC), (ii) ex-interim individually rational[9] (IR), (iii) weakly budget balanced[10] (WBB), and (iv) ex-post efficient, if the intersection of the supports of F and G is an interval with nonzero measure. Thus, the literature is typically focused on identifying the best-possible approximation factor to the first-best GFT of any BNIC, IR, and WBB mechanism [15, 138].

Bilateral Trade with a Broker. In our setup, in addition to the buyer and the seller, there is a broker who governs the entire trade, through whom the buyer and seller exclusively interact. The broker is not interested in maximizing the GFT but instead tries to maximize her own profit by buying from the seller at a lower price and selling to the buyer at a higher price. Following [14], we assume that the broker can be a net source or sink of monetary transfer but that the broker cannot own the item herself. Alternatively, the broker realizes the value 0 for the item and thus does not want to hold on to the item.

Essentially, the broker's *direct* mechanism $\textsc{Mec} = (x, p)$ is defined by an allocation function

[9] A mechanism is ex-interim IR if all the participants are not worse off by participating the mechanism when they know their own valuations but not the others' valuations.

[10] A mechanism is WBB if the payment from the buyer is always at least the payment to the seller in an ex-post manner.

$x : \mathbb{R}^2_{\geq 0} \to [0,1]$ that maps the reported valuations of the buyer and the seller to the probability that the item is transferred from the seller to buyer and a payment function $p : \mathbb{R}^2_{\geq 0} \to \mathbb{R}^2_{\geq 0}$ that maps the reported valuations to the expected payment $p = (p_s, p_b)$, where p_s is the expected payment from the broker to the seller and p_b the expected payment from the buyer to the broker.[11] Note here that, assuming truthful reporting from the buyer and the seller, the broker's expected net profit can be defined as

$$\textsc{Profit} = \int_{a_2}^{b_2} \int_{a_1}^{b_1} (p_b(v,c) - p_s(v,c))\, dF(v)\, dG(c).$$

Similarly, assuming truthful reporting, the buyer's expected utility if his true valuation is v can be written as

$$u_b(v) = \int_{a_2}^{b_2} (x(v,c) \cdot v - p_b(v,c))\, dG(c),$$

and the seller's expected utility if his true valuation is c can be written as

$$u_s(c) = \int_{a_1}^{b_1} (p_s(v,c) - x(v,c) \cdot c)\, dF(v).$$

Desiderata. We are interested in designing an *ex-interim individually-rational* mechanism so that each agent's ex-interim payoff after observing his own valuation without knowing the other's is always nonnegative, *i.e.,* such that

$$u_b(v) \geq 0, u_s(c) \geq 0$$

[11] Note that we consider expected payments here, as the mechanism may randomize the payments.

for any v and c.

A broker's mechanism is *Bayesian-Nash incentive-compatible* (BNIC) if it satisfies

$$\mathbb{E}_{c\sim G}\left[x(v,c)\cdot v - p_b(v,c)\right] \geq \mathbb{E}_{c\sim G}\left[x(\tilde{v},c)\cdot v - p_b(\tilde{v},c)\right]$$

for all v and $\tilde{v}$ and

$$\mathbb{E}_{v\sim F}\left[p_s(v,c) - x(v,c)\cdot c\right] \geq \mathbb{E}_{v\sim F}\left[p_s(v,\tilde{c}) - x(v,\tilde{c})\cdot c\right]$$

for all c and $\tilde{c}$. The first inequality corresponds to the buyer's ex-interim incentive to truthfully report his private value, and the second inequality corresponds to the seller's ex-interim incentive to truthfully report his private value. By the revelation principle [128], any BNIC mechanism can be simulated by a BNIC direct mechanism, so we consider BNIC direct mechanisms without loss of generality.

A broker's mechanism is *dominant-strategy incentive-compatible* (DSIC) if it satisfies

$$x(v,c)\cdot v - p_b(v,c) \geq x(\tilde{v},c)\cdot v - p_b(\tilde{v},c)$$

for every $v, \tilde{v}$, and c and

$$p_s(v,c) - x(v,c)\cdot c \geq p_s(v,\tilde{c}) - x(v,\tilde{c})\cdot c,$$

for every $c, \tilde{c}$, and v. This is a strictly stronger notion than BNIC, *i.e.,* DSIC implies BNIC, since each player is always at least as well off by reporting his true valuation regardless of the other player's valuation.

In this paper, we seek to find the approximation factor to the first-best GFT for different classes of mechanisms. Formally, a mechanism is an α-approximation to the first-best GFT if it satisfies GFT $\geq \alpha$FB-GFT.

Distribution Types. Finally, we introduce several classes of distributions that will be useful in presenting our results. We say that the buyer's distribution is regular if $\phi_F(x) = x - (1 - F(x))/f(x)$ is nondecreasing, and we say that the seller's distribution is regular if $\phi_G(x) = x + G(x)/g(x)$ is nondecreasing. Note that this is a two-sided-market version of the standard regularity condition in single-sided auctions ([128]), where ϕ_F and ϕ_G are usually referred to as the virtual valuations of the buyer and seller, respectively.[12] The hazard rates h_b and h_s of the buyer's and seller's distributions F and G are defined respectively as $h_b(x) = f(x)/(1 - F(x))$ and $h_s(x) = g(x)/G(x)$. We say that the buyer's (resp. seller's) distribution has monotone hazard rate (MHR) if $h_b(x)$ (resp. $h_s(x)$) is nondecreasing (resp. nonincreasing) for $x \geq 0$. It is straightforward to check that MHR implies regularity, but not the other way around. The family of MHR distributions includes the well-known Gaussian, exponential, and uniform distributions whose tails can be no heavier than an exponential distribution with hazard rate 0. However, regular distributions can be subsumed under a broader class of distributions with fatter tails, *e.g.,* $f(x) = x^{-(1+\alpha)}$ for some $\alpha \geq 1$. Lastly, we often consider a symmetric setting in which the buyer's and seller's distributions are equivalent, *i.e.,* where $F = G$. Throughout our paper, we always assume that the expectations of F and G are finite.

[12] In general, even if the density function does not exist, one can define the buyer's distribution to be regular if $\ln(1 - F(x))$ is concave, and analogously for the seller.

6.3 Posted-Pricing Mechanisms

We consider here a broker who implements a posted-pricing (fixed-price) mechanism, *i.e.,* she announces fixed prices to the buyer and seller, respectively. Formally, a posted-pricing mechanism in our setting consists of two fixed prices (p, q). The broker simultaneously offers to buy the item from the seller at price q and sell the item to the buyer at price p. Since the broker strategizes to maximize her profit, it follows that $p \geq q$. After the broker makes her offers, the buyer and seller each decide whether to accept their respective offer in a *take-it-or-leave* manner. If both agree to their offers, then the trade occurs and the item is transferred from the seller to the buyer. Otherwise, no trade occurs.

Under a posted-pricing mechanism with prices (p, q), the gains-from-trade and the broker's expected profit can be written as

$$\text{GFT} = \int_0^q \int_p^\infty (v - c)\, dF(v)\, dG(c).$$
$$\textsc{Profit} = \int_0^q \int_p^\infty (p - q)\, dF(v)\, dG(c).$$

From the above definition, it is clear that any posted-pricing mechanism that maximizes the expected profit for the broker satisfies

$$\begin{aligned}(p^*, q^*) &= \operatorname*{argmax}_{(p,q)\in\mathbb{R}^2_{\geq 0}} \int_0^q \int_p^\infty (p - q)\, dF(v)\, dG(c) \\ &= \operatorname*{argmax}_{(p,q)\in\mathbb{R}^2_{\geq 0}} (p - q)G(q)(1 - F(p)).\end{aligned}$$

We seek to analyze the GFT of any such optimal mechanism compared to the first-best GFT.

We first provide the proof of Theorem 6.1.1.

Proof of Theorem 6.1.1. Observe first that

$$\begin{aligned}
\text{GFT} &= \int_p^\infty \int_0^q (v-c)\,dG(c)\,dF(v) \\
&= \int_p^\infty \int_0^q (v-p+p-q+q-c)\,dG(c)\,dF(v) \\
&= (p-q)(1-F(p))G(q) + \int_p^\infty \int_0^q (v-p)\,dG(c)\,dF(v) + \int_p^\infty \int_0^q (q-c)\,dG(c)\,dF(v). \\
&= \text{PROFIT} + G(q)\int_p^\infty (v-p)\,dF(v) + (1-F(p))\int_0^q (q-c)\,dG(c).
\end{aligned}$$

Integrating by parts, we have that

$$\begin{aligned}
\int_p^\infty (v-p)\,dF(v) &= \Big[-(v-p)(1-F(v))\Big]_p^\infty + \int_p^\infty (1-F(v))\,dv \\
&= \int_p^\infty (1-F(v))\,dv,
\end{aligned}$$

where we use $1-F(v)$ instead of simply $F(v)$ for the integration-by-parts to make the first term exactly zero.[13]

Similarly, for the seller's integral, we have that

$$\begin{aligned}
\int_0^q (q-c)\,dG(c) &= \Big[(q-c)G(c)\Big]_0^q + \int_0^q G(c)\,dc \\
&= \int_0^q G(c)\,dc.
\end{aligned}$$

[13] Since the expectation of F is finite, we have that $v(1-F(v)) = v\int_v^\infty dF(x) \leq \int_v^\infty x\,dF(x)$ and $\lim_{v\to\infty}\left(\int_v^\infty x\,dF(x)\right) = \mathbb{E}_{v\sim F}[v] - \lim_{v\to\infty}\left(\int_0^v x\,dF(x)\right) = 0$, so $\lim_{v\to\infty} v(1-F(v)) = 0$.

Hence, we have that

$$\begin{aligned}
\text{GFT} &= \textsc{Profit} + G(q)\int_p^\infty (1-F(x))\,dx + (1-F(p))\int_0^q G(x)\,dx \\
&\geq \textsc{Profit} + \min(G(q), 1-F(p))\left(\int_p^\infty (1-F(x))\,dx + \int_0^q G(x)\,dx\right) \\
&\geq \textsc{Profit} + \min(G(q), 1-F(p))\left(\int_p^\infty (1-F(x))G(x)\,dx + \int_0^q G(x)(1-F(x))\,dx\right) \\
&= \textsc{Profit} + \min(G(q), 1-F(p))\left(\int_0^\infty (1-F(x))G(x)\,dx - \int_q^p (1-F(x))G(x)\,dx\right).
\end{aligned}$$

Finally, to complete the proof, we use the following equality proven by [120], which we reiterate here for completeness:

$$\begin{aligned}
\text{FB-GFT} &= \int_0^\infty \int_0^v (v-c)\,dG(c)\,dF(v) \\
&= \int_0^\infty \left(\Big[(v-c)G(c)\Big]_0^v + \int_0^v G(c)\,dc\right) dF(v) \\
&= \int_0^\infty \int_0^v G(c)\,dc\,dF(v) \\
&= \Big[-(1-F(v))\int_0^v G(c)dc\Big]_0^\infty + \int_0^\infty (1-F(v))G(v)\,dv \\
&= \int_0^\infty (1-F(v))G(v)\,dv, \qquad (6.2)
\end{aligned}$$

where the limit converges to zero since $(1-F(v))\int_0^v G(c)\,dc \leq v(1-F(v))$ and $\lim_{v\to\infty} v(1-F(v)) = 0$ (see footnote above). This finishes the proof. a

Optimal Prices: Symmetric Agents, MHR Distributions. We here provide the proof of Theorem 6.1.3. As we discussed in Section 6.1.1, the proof mainly follows from proving the existence of quantiles $\mu_F(\alpha)$ and $\mu_F(\beta)$ such that $1-\alpha$ and β are absolute constants regardless of the distributions.

Proof of Theorem 6.1.3. Recall that for a cumulative distribution function F and a constant $0 \leq \alpha \leq 1$, we use $\mu_F(\alpha)$ to denote the α-quantile of F, *i.e.,* $\mu_F(\alpha) = \inf\{t : F(t) \geq \alpha\}$. Now, let $0 \leq \beta < \alpha \leq 1$ be constants to be determined later. Furthermore, suppose that

$$\int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1 - F(x))F(x)\,dx \geq \frac{1}{C}\textsc{FB-GFT}$$

for some constant $C \geq 1$ to be determined later. Now, note that

$$\begin{aligned}\textsc{GFT}(p^*, q^*) &\geq \textsc{Profit}(p^*, q^*) \geq \textsc{Profit}(\mu_F(\alpha), \mu_F(\beta)) \\ &\geq (\mu_F(\alpha) - \mu_F(\beta))\beta(1 - \alpha),\end{aligned}$$

where (p^*, q^*) are an arbitrary pair of prices that maximizes the expected profit of the broker. Therefore, we have that

$$\textsc{GFT}(p^*, q^*) \geq (\mu_F(\alpha) - \mu_F(\beta))\beta(1 - \alpha) \geq \beta(1 - \alpha) \int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1 - F(x))F(x)\,dx \geq \frac{\beta(1 - \alpha)}{C}\textsc{FB-GFT}.$$

Thus, it remains to find constants α, β, C such that $\int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1 - F(x))F(x)dx \geq \frac{1}{C}\textsc{FB-GFT}$. It suffices to find constants α, β, C such that

$$(C - 1) \int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1 - F(x))F(x)\,dx \geq \int_0^{\mu_F(\beta)} (1 - F(x))F(x)\,dx + \int_{\mu_F(\alpha)}^{\infty} (1 - F(x))F(x)\,dx, \tag{6.3}$$

since by (6.2), we have that

$$\begin{aligned}\text{FB-GFT} &= \int_0^\infty (1-F(x))F(x)\,dx \\ &= \int_0^{\mu_F(\beta)} (1-F(x))F(x) + \int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1-F(x))F(x) + \int_{\mu_F(\alpha)}^\infty (1-F(x))F(x).\end{aligned}$$

Now, since we are considering the symmetric setting with MHR distributions, we know that $(1-F(x))/f(x)$ is nonincreasing and that $F(x)/f(x)$ is nondecreasing. Hence, for the integral on the left-hand side of (6.3), we have that

$$\begin{aligned}\int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1-F(x))F(x)\,dx &\geq \int_{\mu_F(\beta)}^{\mu_F(\alpha)} \frac{1-F(x)}{f(x)} f(x)F(x)\,dx \geq \frac{1-\alpha}{f(\mu_F(\alpha))} \int_{\mu_F(\beta)}^{\mu_F(\alpha)} F(x)\,dF(x) \\ \int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1-F(x))F(x)\,dx &\geq \int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1-F(x))\frac{F(x)}{f(x)} f(x)\,dx \geq \frac{\beta}{f(\mu_F(\beta))} \int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1-F(x))\,dF(x).\end{aligned}$$

Therefore, we have that

$$\begin{aligned}\int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1-F(x))F(x)\,dx &\geq \frac{1}{2}\left(\frac{1-\alpha}{f(\mu_F(\alpha))} \int_{\mu_F(\beta)}^{\mu_F(\alpha)} F(x)\,dF(x) + \frac{\beta}{f(\mu_F(\beta))} \int_{\mu_F(\beta)}^{\mu_F(\alpha)} (1-F(x))\,dF(x)\right) \\ &= \frac{1}{2}\left(\frac{1-\alpha}{f(\mu_F(\alpha))} \cdot \frac{\alpha^2-\beta^2}{2} + \frac{\beta}{f(\mu_F(\beta))}\left(\alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right)\right) \qquad (6.4) \\ &\geq \frac{1}{2}\min\left(\frac{\alpha^2-\beta^2}{2}, \alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right)\left(\frac{1-\alpha}{f(\mu_F(\alpha))} + \frac{\beta}{f(\mu_F(\beta))}\right). \qquad (6.5)\end{aligned}$$

Now, consider the right-hand side in (6.3). The first term can be upper bounded by

$$\begin{aligned}\int_0^{\mu_F(\beta)} (1-F(x))F(x)\,dx &= \int_0^{\mu_F(\beta)} \frac{F(x)}{f(x)}(1-F(x))f(x)\,dx \\ &\le \frac{\beta}{f(\mu_F(\beta))}\int_0^{\mu_F(\beta)} (1-F(x))\,dF(x) \\ &= \frac{\beta}{f(\mu_F(\alpha))}\left(\beta - \frac{\beta^2}{2}\right). \end{aligned} \tag{6.6}$$

Similarly, the second term can be bounded above by

$$\begin{aligned}\int_{\mu(\alpha)}^{\infty} (1-F(x))F(x)\,dx &= \int_{\mu(\alpha)}^{\infty} \frac{1-F(x)}{f(x)}F(x)f(x)\,dx \\ &\le \frac{1-\alpha}{f(\mu_F(\alpha))}\int_{\mu(\alpha)}^{1} F(x)\,dF(x) \\ &= \frac{1-\alpha}{f(\mu_F(\alpha))}\left(\frac{1-\alpha^2}{2}\right). \end{aligned} \tag{6.7}$$

Thus, by (6.6) and (6.7), the right-hand side in (6.3) can be upper-bounded by

$$\int_0^{\mu_F(\beta)} (1-F(x))F(x)\,dx + \int_{\mu(\alpha)}^{\infty} (1-F(x))F(x)\,dx \le \max\left(\beta - \frac{\beta^2}{2}, \frac{1-\alpha^2}{2}\right)\left(\frac{1-\alpha}{f(\mu_F(\alpha))} + \frac{\beta}{f(\mu_F(\beta))}\right). \tag{6.8}$$

Hence, from (6.3), we see that it suffices by (6.5) and (6.8) to find constants α, β, C such that

$$\frac{C-1}{2}\min\left(\frac{\alpha^2-\beta^2}{2}, \alpha - \beta - \frac{\alpha^2-\beta^2}{2}\right) \ge \max\left(\beta - \frac{\beta^2}{2}, \frac{1-\alpha^2}{2}\right).$$

By (roughly) numerically optimizing the parameters α, β and C here, taking $\alpha = 2/3, \beta = 1/3$, and

$C = 13/3$, we see that

$$\text{GFT}(p^*, q^*) \geq \frac{\beta(1-\alpha)}{C}\text{FB-GFT} = \frac{1}{39}\text{FB-GFT}.$$

a

Optimal Prices: Bounding via Hazard Rates. We provide the proof of Theorem 6.1.4 here. The proof starts with the observation that FB-GFT can be written in terms of the integral of $|v - c|$. Then, noting that the optimal broker's profit maximizes $(p - q)(1 - F(p))G(q)$, we lower bound the profit by an averaging argument over a desirable region. Combining this with the bounded hazard rates completes the proof. For simplicity, we restrict the support of the distributions here to be $[0, 1]$ without loss of generality, *e.g.,* see Remark 2 in [15].

Proof of Theorem 6.1.4. Note first that

$$\begin{aligned}
\int_0^1 \int_v^1 |v - c|\, dF(c)\, dF(v) &= \int_0^1 \int_0^c |v - c|\, dF(v)\, dF(c) \\
&= \int_0^1 \int_0^v |c - v|\, dF(c)\, dF(v) \\
&= \int_0^1 \int_0^v |v - c|\, dF(c)\, dF(v) \\
&= \text{FB-GFT}.
\end{aligned}$$

Thus, we have that

$$\text{FB-GFT} = \frac{1}{2}\int_0^1 \int_0^1 |v - c|\, dF(c)\, dF(v).$$

On the other hand, since the broker uses an optimal pair of prices (p^*, q^*), we have that

$$\textsc{Profit} = (p^* - q^*)(1 - F(p^*))F(q^*) \geq (p - q)(1 - F(p))F(q),$$

for any $0 \leq q \leq p \leq 1$. Integrating both sides over the region of $\mathbb{R}^2_{\geq 0}$ that satisfies $p \geq q$, we obtain

$$\begin{aligned}
\frac{\textsc{Profit}}{2} &\geq \int_0^1 \int_q^1 (p - q)(1 - F(p))F(q)\, dp\, dq \\
&= \int_0^1 \int_q^1 (p - q)\frac{1 - F(p)}{f(p)}\frac{F(q)}{f(q)} f(p)f(q)\, dp\, dq. \\
&\geq \int_0^1 \int_q^1 (p - q)\frac{1}{M^2} f(p)f(q)\, dp\, dq \qquad (6.9) \\
&= \frac{1}{M^2}\int_0^1 \int_q^1 (p - q) f(p)f(q)\, dp\, dq \\
&= \frac{1}{M^2}\int_0^1 \int_q^1 (p - q)\, dF(p)\, dF(q), \qquad (6.10)
\end{aligned}$$

where in the last inequality, we use the fact that both hazard rates are uniformly bounded above by M and thus their inverses are uniformly bounded below by $1/M$.

Analogously, by expanding the integral from the seller's side first, we have

$$\begin{aligned}
\frac{\textsc{Profit}}{2} &\geq \int_0^1 \int_0^p (p - q)(1 - F(p))F(q)\, dq\, dp \\
&= \int_0^1 \int_0^p (p - q)\frac{1 - F(p)}{f(p)}\frac{F(q)}{f(q)} f(p)f(q)\, dq\, dp \\
&\geq \frac{1}{M^2}\int_0^1 \int_0^p (p - q)\, dF(q)\, dF(p) \\
&= \frac{1}{M^2}\int_0^1 \int_0^q (q - p)\, dF(p)\, dF(q).
\end{aligned}$$

Combining this with (6.10), we obtain

$$\begin{aligned}\textsc{Profit} &\geq \frac{1}{M^2}\int_0^1\int_0^1 |p-q|\, dF(p)dF(q)\\ &= \frac{2}{M^2}\text{FB-GFT},\end{aligned}$$

and the proof follows from the fact that GFT $\geq$ PROFIT.

a

It is worth noting that the approximation factor here cannot be better than $1/2$, since we must have that $M \geq 2$ according to the problem setup. To see why, observe that

$$\frac{1-F(x)}{f(x)} \geq \frac{1}{M}, \frac{F(x)}{f(x)} \geq \frac{1}{M}$$

implies that $M \geq 2$ as follows:

$$\frac{1}{M} \leq \frac{F(x)}{f(x)} = 1 - \frac{1-F(x)}{f(x)} \leq 1 - \frac{1}{M}.$$

Quantile-based Prices. As mentioned in Section 6.1.1, the decomposition in Theorem 6.1.1 cannot be directly applied to obtain a constant approximation factor for the first-best GFT in general, as it depends on an integral involving the buyer's and seller's CDFs in the interval between the two posted prices. This brings us to the question of whether we can obtain an approximation factor purely characterized by the quantile probabilities at the offered prices, as suggested by Theorem 6.1.5.

We formally prove Theorem 6.1.5 in what follows:

Proof of Theorem 6.1.5. For any $0 \leq c \leq \mu_G(\beta) \leq \mu_F(\alpha)$,

$$\begin{aligned}\int_c^\infty (v-c)\,dF(v) &= \int_c^{\mu_F(\alpha)} (v-c)\,dF(v) + \int_{\mu_F(\alpha)}^\infty (v-c)\,dF(v) \\ &\leq \frac{\alpha}{1-\alpha}\int_{\mu_F(\alpha)}^\infty (v-c)\,dF(v) + \int_{\mu_F(\alpha)}^\infty (v-c)\,dF(v) \\ &= \frac{1}{1-\alpha}\int_{\mu_F(\alpha)}^\infty (v-c)\,dF(v), \end{aligned} \tag{6.11}$$

where we used the fact that

$$\int_c^{\mu_F(\alpha)} (v-c)\,dF(v) \leq \frac{\alpha}{1-\alpha}\int_{\mu_F(\alpha)}^\infty (v-c)\,dF(v).$$

Observe that this holds because (1) the probability mass of the first integral is at most α while the probability mass of the second integral is exactly $1-\alpha$ and (2) the value of $v-c$ at any point in the probability mass of the second integral is always as least as much as the value of $v-c$ at any point in the probability mass of the first integral. In fact, reusing this idea, we have that

$$\begin{aligned}\int_{\mu_G(\beta)}^\infty \int_c^\infty (v-c)\,dF(v)\,dG(c) &\leq \frac{1-\beta}{\beta}\int_0^{\mu_G(\beta)} \int_c^\infty (v-c)\,dF(v)\,dG(c) \\ \int_0^\infty \int_c^\infty (v-c)\,dF(v)\,dG(c) &\leq \frac{1}{\beta}\int_0^{\mu_G(\beta)} \int_c^\infty (v-c)\,dF(v)\,dG(c). \end{aligned} \tag{6.12}$$

Finally, combining (6.11) and (6.12), we obtain

$$\int_0^{\mu_G(\beta)} \int_{\mu_F(\alpha)}^\infty (v-c)\,dF(v)\,dG(c) \geq \beta(1-\alpha)\int_0^\infty \int_c^\infty (v-c)\,dF(v)\,dG(c),$$

as desired. a

Proof of Corollary 6.1.6. Note that drawing samples $p \sim F$ and $q \sim G$ is equivalent to sampling quantiles α and β of F and G, respectively, uniformly at random over $[0, 1]$. Thus, letting $\text{GFT}(x, y)$ denote the gains-from-trade of the posted-pricing mechanism (x, y), we have that

$$\begin{aligned}
&\mathbb{E}_{p\sim F, q\sim G}\left[\text{GFT}(\max(p,q), \min(p,q))\right] \\
&\quad = \int_0^1 \int_\beta^1 \text{GFT}(\mu_F(\alpha), \mu_G(\beta))\, d\alpha\, d\beta + \int_0^1 \int_0^\beta \text{GFT}(\mu_G(\beta), \mu_F(\alpha))\, d\alpha\, d\beta \\
&\quad \geq \int_0^1 \int_\beta^1 \beta(1-\alpha)\text{FB-GFT}\, d\alpha\, d\beta + \int_0^1 \int_0^\beta \alpha(1-\beta)\text{FB-GFT}\, d\alpha\, d\beta = \frac{1}{12}\text{FB-GFT},
\end{aligned}$$

where the inequality holds by Theorem 6.1.5, since $F = G$. a

We conclude by remarking that having access to more samples may not necessarily increase the approximation factor to the first-best gains-from-trade, since it may be more optimal for the better-informed broker to set prices that further restrict trade.

6.4 Optimal Mechanisms

We now consider the case where the broker runs any general BNIC and IR mechanism that maximizes her profit. We assume here that the supports of F and G are $[a_1, b_1]$ and $[a_2, b_2]$, respectively, *i.e.,* both are bounded, following [14].[14] By Theorem 6.1.7, the broker's optimal mechanism exhibits

[14]Indeed, if we consider F and G with unbounded supports, we observe that the characterization of [14] of an optimal BNIC and IR mechanism does not carry over, which is elaborated on more in the proof of Theorem 6.1.9.

the following GFT, assuming regular distributions:

$$\begin{aligned}\text{GFT}_{F,G} &= \int_{a_2}^{b_2}\int_{a_1}^{b_1}(v-c)\mathbf{1}\left\{\phi_F(v)\geq\phi_G(c)\right\}\,dF(v)\,dG(c)\\ &= \int_{a_2}^{b_2}\int_{a_1}^{b_1}(v-c)\mathbf{1}\left\{v-c\geq\frac{1-F(v)}{F'(v)}+\frac{G(c)}{G'(c)}\right\}\,dF(v)\,dG(c).\end{aligned}$$

As usual, the first-best GFT is given by

$$\text{FB-GFT}_{F,G} = \int_{a_2}^{b_2}\int_{a_1}^{b_1}(v-c)\mathbf{1}\left\{v-c\geq 0\right\}\,dF(v)\,dG(c).$$

Uniform distributions. To prove Theorem 6.1.8, we require a technical lemma that will simplify our analysis. The lemma states that the factor to which the broker's optimal mechanism approximates the first-best GFT is invariant to stretches and shifts in F and G, provided that we perform the same stretches and shifts to F and G.

Lemma 6.4.1. *Let F and G be regular distributions supported on $[a_1,b_1]$ and $[a_2,b_2]$, respectively. Define $F^*(v)=F((v-k_1)/k_2)$ and $G^*(c)=G((c-k_1)/k_2)$ for $k_1\in\mathbb{R}$ and $k_2\in\mathbb{R}_{>0}$. If $\text{GFT}_{F,G}=\alpha\text{FB-GFT}_{F,G}$, then $\text{GFT}_{F^*,G^*}=\alpha\text{FB-GFT}_{F^*,G^*}$.*

Proof. We first prove that F^* and G^* are regular. Note that to show that $\phi_{F^*}(v^*)$ is nondecreasing on its support $[k_2a_1+k_1,k_2b_1+k_1]$, it suffices to show that $\phi_{F^*}(k_2v+k_1)$ is nondecreasing on $[a_1,b_1]$.

Indeed, since

$$\begin{aligned}\phi_{F^*}(k_2 v + k_1) &= (k_2 v + k_1) - \frac{1 - F^*(k_2 v + k_1)}{(F^*)'(k_2 v + k_1)} \\ &= (k_2 v + k_1) - \frac{1 - F(v)}{F'(v)/k_2} \\ &= k_1 + k_2 \phi_F(v)\end{aligned}$$

and $\phi_F(v)$ is nondecreasing, F^* is regular. A similar argument will show that G^* is regular.

Now, we have that

$$\begin{aligned}\text{GFT}_{F,G} &= \int_{a_2}^{b_2} \int_{a_1}^{b_1} (v - c)\mathbf{1}\left\{v - c \geq \frac{1 - F(v)}{F'(v)} + \frac{G(c)}{G'(c)}\right\} F'(v)\, dv\; G'(c)\, dc \\ &= \int_{a_2^*}^{b_2^*} \int_{a_1^*}^{b_1^*} \frac{v^* - c^*}{k_2}\mathbf{1}\left\{\frac{v^* - c^*}{k_2} \geq \frac{1 - F(\frac{v^* - k_1}{k_2})}{F'(\frac{v^* - k_1}{k_2})} + \frac{G(\frac{c^* - k_1}{k_2})}{G'(\frac{c^* - k_1}{k_2})}\right\} \\ &\quad \cdot \frac{1}{k_2} F'\left(\frac{v^* - k_1}{k_2}\right) dv^*\; \frac{1}{k_2}\, G'\left(\frac{c^* - k_1}{k_2}\right) dc^*,\end{aligned}$$

where we use the change of variables with $v^* = k_2 v + k_1$ and $c^* = k_2 c + k_1$. This can further be expanded as

$$\begin{aligned}\text{GFT}_{F,G} &= \int_{a_2^*}^{b_2^*} \int_{a_1^*}^{b_1^*} \frac{v^* - c^*}{k_2}\mathbf{1}\left\{\frac{v^* - c^*}{k_2} \geq \frac{1 - F^*(v^*)}{k_2 (F^*)'(v^*)} + \frac{G^*(c^*)}{k_2 (G^*)'(c^*)}\right\} (F^*)'(v^*)\, dv^*\; (G^*)'(c^*)\, dc^* \\ &= \int_{a_2^*}^{b_2^*} \int_{a_1^*}^{b_1^*} \frac{v^* - c^*}{k_2}\mathbf{1}\left\{v^* - c^* \geq \frac{1 - F^*(v^*)}{(F^*)'(v^*)} + \frac{G^*(c^*)}{(G^*)'(c^*)}\right\} (F^*)'(v^*)\, dv^*\; (G^*)'(c^*)\, dc^* \\ &= \frac{1}{k_2}\text{GFT}_{F^*,G^*},\end{aligned}$$

where the last equality follows from F^* and G^* being regular distributions.

A similar argument will show that

$$\text{FB-GFT}_{F,G} = \frac{1}{k_2}\text{FB-GFT}_{F^*,G^*}.$$

Thus, if $\text{GFT}_{F,G} = \alpha\text{FB-GFT}_{F,G}$, then

$$\text{GFT}_{F^*,G^*} = k_2\text{GFT}_{F,G} = \alpha k_2\text{FB-GFT}_{F,G} = \alpha\text{FB-GFT}_{F^*,G^*}.$$

a

The proof of Theorem 6.1.8 is deferred to Appendix E.1, as it requires much case work that includes cumbersome algebraic manipulations.

Inapproximability. We provide the proof of Theorem 6.1.9 here, which states that the extent to which the broker's optimal strategic behavior degrades the GFT can be significant.

Proof of Theorem 6.1.9. Recall that by Theorem 6.1.7, the gains-from-trade for any BNIC and IR mechanism that maximizes the broker's expected profit, assuming regular distributions, is given by

$$\mathsf{GFT} = \iint (v-c)\mathbf{1}\left\{v - c \geq \frac{1-F(v)}{F'(v)} + \frac{G(c)}{G'(c)}\right\} dF(v)\, dG(c).$$

Thus, to construct F and G for which such a mechanism achieves a 0-approximation to the gains-from-trade, we want to construct F such that

$$v = \frac{1-F(v)}{F'(v)},$$

for all v, in which case we have that

$$v - c < \frac{1 - F(v)}{F'(v)} + \frac{G(c)}{G'(c)}$$

for all v and c. Solving the above differential equation, we obtain the general solution

$$F(v) = -\frac{a}{v} + 1$$

on $[a, \infty)$ for $a > 0$. However, since the result about the optimal broker's mechanism by [14] only applies to distributions with bounded supports, we now modify F by first truncating it and then extending it using a tangent line.[15] In particular, we define

$$F(v) = \begin{cases} -\frac{a}{v} + 1 & \text{if } a \leq v \leq b \\ \left(-\frac{a}{b} + 1\right) + \frac{a}{b^2}(v - b) & \text{if } b < v \leq 2b \end{cases}$$

for $b > a \geq 1$. Note that F is regular, since

$$\phi_F(v) = \begin{cases} 0 & \text{if } a \leq v \leq b \\ 2v - 2b & \text{if } b < v \leq 2b \end{cases}$$

is increasing. Finally, let G be the uniform (and thus regular) distribution on $[0, 1]$.

We will now show that for any $\epsilon > 0$, there exists $b > a$ such that any BNIC and IR mechanism that maximizes the broker's expected profit induces $\mathsf{GFT} < \epsilon\mathsf{FB\text{-}GFT}$.

[15] Indeed, with the original F, trade never occurs by construction, so $\textsc{Profit} = 0$. This implies that the characterization by [14] does not apply as the broker can simply use a posted-pricing mechanism to obtain a positive profit.

By the construction of F, we have that $v - c < \dfrac{1 - F(v)}{F'(v)} + \dfrac{G(c)}{G'(c)}$ for all $v \in (a, b)$, so

$$\begin{aligned}
\text{GFT} &= \frac{a}{b^2} \int_b^{2b} \int_0^1 (v - c) \mathbf{1} \left\{ v - c \geq \frac{1 - F(v)}{F'(v)} + \frac{G(c)}{G'(c)} \right\} dc\, dv \\
&\leq \frac{a}{b^2} \int_b^{2b} \int_0^1 (v - c)\, dc\, dv \\
&= \frac{a}{b^2} \int_b^{2b} \left(v - \frac{1}{2} \right) dv \\
&= \frac{a}{b^2} \left(\frac{3b^2}{2} - \frac{b}{2} \right) \\
&= \frac{a(3b - 1)}{2b}.
\end{aligned}$$

In addition, by the above derivation,

$$\begin{aligned}
\text{FB-GFT} &= a \int_a^b \int_0^1 (v - c)\, dc\, \frac{1}{v^2}\, dv + \frac{a}{b^2} \int_b^{2b} \int_0^1 (v - c)\, dc\, dv \\
&= a \int_a^b \left[vc - \frac{c^2}{2} \right]_0^1 \frac{1}{v^2} dv + \frac{a(3b - 1)}{2b} \\
&= a \int_a^b \left(v - \frac{1}{2} \right) \frac{1}{v^2}\, dv + \frac{a(3b - 1)}{2b} \\
&= a \left(\ln \left(\frac{b}{a} \right) + \frac{1}{2b} - \frac{1}{2a} \right) + \frac{a(3b - 1)}{2b}.
\end{aligned}$$

Hence,

$$\frac{\text{GFT}}{\text{FB-GFT}} \leq \frac{a(3b - 1)}{2b} \Bigg/ \left(a \left(\ln \left(\frac{b}{a} \right) + \frac{1}{2b} - \frac{1}{2a} \right) + \frac{a(3b - 1)}{2b} \right),$$

which approaches 0 as $b \to \infty$, so $\dfrac{\text{GFT}}{\text{FB-GFT}}$ approaches 0 as $b \to \infty$. a

Inapproximability with a Public Agent. Here, we formally prove Theorem 6.1.10, first for the

public-seller setting and then for the public-buyer setting. To prove the theorem for the public-seller setting, we first introduce some notation:

Definition 6.4.2. Given a direct mechanism MEC $= (x(v), p(v))$[16], we define the following functions for the expected utilities of the buyer, intermediary, and seller, respectively, assuming truthful reporting by the buyer:

- $u_b(v) = x(v) \cdot v - p_b(v)$
- $u_i = \int_a^b (p_b(v) - p_s(v))\, F'(v)\, dv$
- $u_s = \int_a^b (p_s(v) - x(v) \cdot c)\, F'(v)\, dv$

Note that with a public seller, MEC is BNIC if $u_b(v) \geq x(\tilde{v}) \cdot v - p_b(\tilde{v})$ for every v and $\tilde{v}$, and MEC is IR if $u_b(v) \geq 0$ for all v and $u_s \geq 0$.

The following theorem is an analogue of Theorem 3 in [14]:

Theorem 6.4.3. *For any BNIC mechanism* MEC *in the public-seller setting,* $u_b(v)$ *is increasing and*

$$u_i + u_b(a) + u_s = \int_a^b \left(\left(v - \frac{1 - F(v)}{F'(v)} \right) - c \right) x(v)\, F'(v)\, dv.$$

Proof. Since MEC is BNIC, for every v and v',

$$u_b(v) = x(v) \cdot v - p_b(v) \geq x(v') \cdot v - p_b(v')$$

and

$$u_b(v') = x(v') \cdot v' - p_b(v') \geq x(v) \cdot v' - p_b(v).$$

[16] In the public-seller setting, the allocation function x and payment function p of MEC only depend on the buyer's reported valuation v.

Thus,

$$(v - v')x(v) \geq u_b(v) - u_b(v') \geq (v - v')x(v').$$

In particular, if $v > v'$, then $x(v) \geq x(v')$, so $x(v)$ is increasing and hence integrable. Therefore, $u_b'(v) = x(v)$ almost everywhere, so

$$u_b(v) = u_b(a) + \int_a^v x(t)\,dt \tag{6.13}$$

and $u_b(\cdot)$ is increasing.

Now, by (6.13), we have that

$$\begin{aligned}
\int_a^b (v - c)x(v)\,F'(v)\,dv - u_i &= \int_a^b (vx(v) - p_b(v))\,F'(v)\,dv + \int_a^b (p_s(v) - cx(v))\,F'(v)\,dv \\
&= \int_a^b u_b(v)F'(v)\,dv + u_s \\
&= \int_a^b \left(u_b(a) + \int_a^v x(t)\,dt\right) F'(v)\,dv + u_s \\
&= u_b(a) + \int_a^b \int_a^v x(t)\,dt\,F'(v)\,dv + u_s \\
&= u_b(a) + \int_a^b \int_t^b F'(v)\,dv\,x(t)\,dt + u_s \\
&= u_b(a) + \int_a^b (1 - F(t))\,x(t)\,dt + u_s.
\end{aligned}$$

Finally, from the first and last expressions in this chain of equations, we obtain

$$u_i + u_b(a) + u_s = \int_a^b \left(\left(v - \frac{1 - F(v)}{F'(v)}\right) - c\right) x(v)\,F'(v)\,dv.$$

a

Next, the following is an analogue of Theorem 4 in [14], tailored to the public-seller setting.

Theorem 6.4.4. *Suppose F is a regular distribution in the public-seller setting. Then, among all BNIC and IR mechanisms, the broker's expected profit is maximized by a mechanism in which the object is transferred to the buyer if and only if $\phi_F(v) \geq c$.*

Proof. By Theorem 6.4.3, we have that

$$u_i = \int_a^b \left(\left(v - \frac{1 - F(v)}{F'(v)} \right) - c \right) x(v)\, F'(v)\, dv - u_b(a) - u_s.$$

Since we want an IR mechanism MEC that maximizes u_i, we want MEC to satisfy the following constraints:

1. $u_b(a) = 0$
2. $u_s = 0$
3. $x(v) = \begin{cases} 1 & \text{if } \phi_F(v) \geq c \\ 0 & \text{otherwise} \end{cases}$

It remains to define $p(v) = (p_s(v), p_b(v))$ such that first two constraints are satisfied. We take

$$p_s(v) = x(v)c$$

and

$$p_b(v) = x(v) \cdot \inf\{t \geq a | \phi_F(t) \geq c\}.$$

Then, since $p_b(a) = ax(a)$ and $p_s(v) - x(v)c = 0$ for all v, we have that $u_b(a) = 0$ and $u_s = 0$, as desired.

Note that MEC is BNIC, since if trade occurs, the buyer pays the lowest valuation he could have reported and still have the trade occur. Then, since $u_b(v)$ is increasing by Theorem 6.4.3 and $u_b(a) = 0$, we have that MEC is IR. a

The next corollary directly follows from the above theorem:

Corollary 6.4.5. *Suppose F is a regular distribution in the public-seller setting. Then, the gains-from-trade from any BNIC and IR mechanism that maximizes the broker's expected profit is*

$$\int_a^b (v-c)\mathbf{1}\left\{v - c \geq \frac{1-F(v)}{F'(v)}\right\} F'(v)\,dv.$$

Using similar ideas, we prove the following theorem which is the public-buyer analogue of Theorem 6.4.4. Its proof is deferred to Appendix E.2 due to its similarity.

Theorem 6.4.6. *Suppose G is a regular distribution in the public-buyer setting. Then, the gains-from-trade from any BNIC and IR mechanism that maximizes the broker's expected profit is*

$$\int_a^b (v-c)\mathbf{1}\left\{v - c \geq \frac{G(c)}{G'(c)}\right\} G'(c)\,dc.$$

We are finally ready to prove Theorem 6.1.10.

Proof of Theorem 6.1.10. We first prove the public-seller case. Let $b > a > 0$. Define

$$F(v) = \begin{cases} -\dfrac{a}{v} + 1 & \text{if } a \leq v \leq b \\ \left(-\dfrac{a}{b} + 1\right) + \dfrac{a}{b^2}(v-b) & \text{if } b < v \leq 2b \end{cases},$$

and let G be the singleton distribution supported on $\{a\}$. Then,

$$\begin{aligned}\text{FB-GFT} &= a\int_a^b (v-a)\cdot\frac{1}{v^2}\,dv + \frac{a}{b^2}\int_b^{2b}(v-a)\,dv \\ &= a\left(\ln\left(\frac{b}{a}\right)+\frac{a}{b}-1\right)+\left(\frac{3a}{2}-\frac{a^2}{b}\right),\end{aligned}$$

and since F is regular, by Corollary 6.4.5 and using the facts that $v = (1-F(v))/F'(v)$ on (a,b) and that $b \le (a+2b)/2 \le 2b$,

$$\begin{aligned}\text{GFT} &= \int_a^{2b}(v-a)\mathbf{1}\left\{v-a \ge \frac{1-F(v)}{F'(v)}\right\}F'(v)\,dv \\ &= \frac{a}{b^2}\int_b^{2b}(v-a)\mathbf{1}\left\{v-a \ge \frac{1-F(v)}{F'(v)}\right\}dv \\ &= \frac{a}{b^2}\int_b^{2b}(v-a)\mathbf{1}\left\{v \ge \frac{a+2b}{2}\right\}dv \\ &= \frac{a}{b^2}\int_{(a+2b)/2}^{2b}(v-a)\,dv \\ &= \frac{3a(a-2b)^2}{8b^2}.\end{aligned}$$

Thus, as $b\to\infty$, we have that $\dfrac{\text{GFT}}{\text{FB-GFT}} \to 0$.

Finally, we prove the public-buyer case. Let $0<\delta<\dfrac{1}{2}$. Define

$$G(c) = \begin{cases}\delta c & \text{if } 0 \le c \le 1 \\ \\ \rho(c) & \text{if } 1 < c \le 2\end{cases},$$

where

$$\rho(c) = k - \frac{a}{c-h}$$

and

$$h = \frac{2-3\delta}{1-2\delta},\ a = \delta(1-h)^2,\ k = \delta(2-h)$$

are chosen so that

$$\rho(1) = \delta,\ \rho'(1) = \delta,\ \rho(2) = 1.$$

In addition, let F be the singleton distribution supported on $\{2\}$. Then,

$$\begin{aligned}
\text{FB-GFT} &= \delta \int_0^1 (2-c)\, dc + a \int_1^2 (2-c) \cdot \frac{1}{(c-h)^2}\, dc \\
&= \frac{3}{2}\delta + a\left(\frac{1}{1-h} - \ln\left(\frac{2-h}{1-h}\right)\right) \\
&= \frac{3}{2}\delta + \frac{\delta(1-\delta)^2}{(2\delta-1)^2}\left(\frac{2\delta-1}{1-\delta} - \ln\left(\frac{\delta}{1-\delta}\right)\right) \\
&\geq \frac{\delta(1-\delta)^2}{(2\delta-1)^2}\left(\frac{2\delta-1}{1-\delta} - \ln\left(\frac{\delta}{1-\delta}\right)\right),
\end{aligned}$$

and since G is regular, by Theorem 6.4.6,

$$\begin{aligned}\text{GFT} &= \int_0^2 (2-c)\mathbf{1}\left\{2 \geq c + \frac{G(c)}{G'(c)}\right\} G'(c)\, dc \\ &= \delta \int_0^1 (2-c)\, dc \\ &= \frac{3}{2}\delta,\end{aligned}$$

since $\phi_G(1) = 1 + \dfrac{G(1)}{G'(1)} = 2$ and $\phi_G(\cdot)$ is strictly increasing. Thus,

$$\frac{\text{FB-GFT}}{\text{GFT}} \geq \frac{2(1-\delta)^2}{3(2\delta-1)^2}\left(\frac{2\delta-1}{1-\delta} - \ln\left(\frac{\delta}{1-\delta}\right)\right),$$

so taking the limit as $\delta \to 0$, we obtain

$$\begin{aligned}&\lim_{\delta\to 0}\left(\frac{2(1-\delta)^2}{3(2\delta-1)^2}\left(\frac{2\delta-1}{1-\delta} - \ln\left(\frac{\delta}{1-\delta}\right)\right)\right) \\ &= \lim_{\delta\to 0}\frac{2(1-\delta)}{3(2\delta-1)} - \lim_{\delta\to 0}\frac{2(1-\delta)^2}{3(2\delta-1)^2}\ln\delta + \lim_{\delta\to 0}\frac{2(1-\delta)^2}{3(2\delta-1)^2}\ln(1-\delta) \\ &= -\frac{2}{3} - \lim_{\delta\to 0}\frac{2(1-\delta)^2}{3(2\delta-1)^2}\ln\delta \\ &= \infty,\end{aligned}$$

which completes the proof. a

Inapproximability with Symmetric Agents. Finally, we provide the proof of Theorem 6.1.11 here. The proof mainly follows from considering the same distribution exploited in Theorem 6.1.9.

Proof. Let $b > a > 0$. Let

$$F(v) = \begin{cases} -\dfrac{a}{v} + 1 & \text{if } a \le v \le b \\ \left(-\dfrac{a}{b} + 1\right) + \dfrac{a}{b^2}(v - b) & \text{if } b < v \le 2b, \end{cases}$$

and let $G = F$. Then,

$$\begin{aligned} \text{FB-GFT} &= a^2 \int_a^b \int_a^v (v - c)\, \frac{1}{c^2}\, dc\, \frac{1}{v^2}\, dv + \frac{a^2}{b^2} \int_b^{2b} \int_a^b (v - c)\, \frac{1}{c^2}\, dc\, dv + \frac{a^2}{b^4} \int_b^{2b} \int_b^v (v - c)\, dc\, dv \\ &= a^2 \int_a^b \left[-\frac{v}{c} - \ln c\right]_a^v \frac{1}{v^2}\, dv + \frac{a^2}{b^2} \int_b^{2b} \left[-\frac{v}{c} - \ln c\right]_a^b dv + \frac{a^2}{b^4} \int_b^{2b} \left[vc - \frac{c^2}{2}\right]_b^v dv \\ &= a^2 \int_a^b \left(\frac{v}{a} - 1 - \ln\left(\frac{v}{a}\right)\right) \frac{1}{v^2}\, dv + \frac{a^2}{b^2} \int_b^{2b} \left(\frac{v}{a} - \frac{v}{b} - \ln\left(\frac{b}{a}\right)\right) dv + \frac{a^2}{2b^4} \int_b^{2b} (v - b)^2\, dv. \end{aligned}$$

Expanding the integral once again, we obtain

$$\text{FB-GFT} \ge a \ln(b) - a \ln a - 2a - \frac{3a^2}{2b}.$$

In addition, by Theorem 6.1.7, we have that

$$\begin{aligned}
\text{GFT} &= \iint (v-c)\mathbf{1}\left\{v - c \geq \frac{1-F(v)}{F'(v)} + \frac{G(c)}{G'(c)}\right\} F'(v)\,dv\,G'(c)\,dc \\
&= \frac{a}{b^2}\int_b^{2b}\int (v-c)\mathbf{1}\left\{v - c \geq \frac{1-F(v)}{F'(v)} + \frac{G(c)}{G'(c)}\right\} G'(c)\,dc\,dv \\
&\leq \frac{a}{b^2}\int_b^{2b}\int (v-c)\mathbf{1}\left\{v - c \geq 0\right\} G'(c)\,dc\,dv \\
&= \frac{a^2}{b^2}\int_b^{2b}\int_a^b (v-c)\frac{1}{c^2}\,dc\,dv + \frac{a^2}{b^4}\int_b^{2b}\int_b^v (v-c)\,dc\,dv \\
&= \frac{3a}{2} - \frac{3a^2}{2b} - \frac{a^2}{b}\ln\left(\frac{b}{a}\right) + \frac{a^2}{6b} \\
&\leq \frac{3a}{2} + \frac{a^2}{6b}.
\end{aligned}$$

Thus, since $\text{GFT} \sim \frac{3a}{2}$ and $\text{FB-GFT} \sim a\ln b$, we have that $\frac{\text{GFT}}{\text{FB-GFT}} \to 0$ as $b \to \infty$. a

Inapproximability of Social Welfare. Interestingly, in several of the settings where we have proven inapproximability of the first-best gains-from-trade, the inapproximability carries over to the first-best social welfare, even though the latter is at least as easy to approximate as the former.

Proof of Theorem 6.1.12. Suppose we have a family of problem instances $\{(F_\delta, G_\delta)\}_{\delta\in\mathbb{R}_+}$ satisfying the following:

1. The supports of the G_δ are uniformly upper-bounded by some constant M.
2. For any (F_δ, G_δ), we have that $\text{GFT} < \delta\text{FB-GFT}$.
3. $\lim_{\delta\to 0}\text{FB-GFT} = \infty$

Then,

$$\frac{\text{SW}}{\text{FB-SW}} = \frac{\mathbb{E}_{c\sim G_\delta}[c] + \text{GFT}}{\mathbb{E}_{c\sim G_\delta}[c] + \text{FB-GFT}} \leq \frac{M + \text{GFT}}{M + \text{FB-GFT}} \leq \frac{M + \delta\text{FB-GFT}}{M + \text{FB-GFT}} = \frac{M/\text{FB-GFT} + \delta}{M/\text{FB-GFT} + 1}$$

for any (F_δ, G_δ), so since $\frac{M/\text{FB-GFT} + \delta}{M/\text{FB-GFT} + 1} \to 0$ as $\delta \to 0$, we have that $\frac{\text{SW}}{\text{FB-SW}} \to 0$ as $\delta \to 0$.

To complete the proof, note that by taking (F_δ, G_δ) to be the problem instance from Theorem 6.1.9 (respectively the public-seller problem instance from Theorem 6.1.10) such that GFT $<$ δFB-GFT, we obtain a family of problem instances $\{(F_\delta, G_\delta)\}_{\delta\in\mathbb{R}_+}$ satisfying the three properties above. a

Chapter 7: Single-sample Bilateral Trade with a Broker

This chapter presents our results on single-sample bilateral trade with a broker [18] (WWW'26).

7.1 Introduction

The bilateral trade problem, introduced in the seminal work of [14], is a fundamental problem in mechanism design in which a buyer wishes to sell a single item to a seller. Both agents hold private valuations on the item and aim to maximize their quasi-linear utility. A mechanism designer's objective is to maximize the social welfare induced by the trade, *i.e.,* transfer the item whenever the buyer's value is larger than the seller's. Such a mechanism is called ex-post efficient. At the same time, a mechanism should satisfy reasonable desiderata to (i) incentivize the agents to participate (individual

rationality) (ii) incentivize the agents to report their true valuations (incentive compatibility) and (iii) neither subsidize the market nor benefit from it (budget-balancedness).

Unfortunately, [14] prove an impossibility result barring the existence of ex-post efficient mechanisms that are simultaneously incentive-compatible (IC), individually-rational (IR), and budget-balanced (BB). This pessimistic result holds even when IC is loosened to hold in an ex-ante manner, known as Bayesian incentive-compatibility (BIC), and BB is relaxed to weak budget-balancedness (WBB), *i.e.,* the mechanism does not run a deficit. As such, there has been a long line of work studying the *approximability* of the first-best efficiency by such mechanisms, *e.g.,* [120], [122], [141], [15], [30].

On the other hand, numerous real-world trades involve two parties exchanging an item through a *broker* due to regulatory policy, market transparency, or privacy issues, *e.g.,* financial markets, foreign exchange, commodity markets, markets for demerit items, or real estate. Importantly, the broker often has objectives misaligned with those of society, *e.g.,* a broker may try to maximize her own profit. In spite of numerous works studying approximate efficiency without a broker,[1] there had been few works studying approximate efficiency with a strategic broker until [148] recently studied the approximability of the first-best gains-from-trade under a profit-maximizing broker who has full knowledge of the buyer's and seller's distributions.

In practice, however, the broker typically does not hold full information on the agents' distributions, but may instead only possess partial information based on previous trades, *e.g.,* a series of bid samples each agent has submitted. Motivated by such scenarios, we consider the approximate efficiency in bilateral trade where the broker has extremely restricted information on the buyer and seller. Formally, we consider a broker who only has access to a single sample from each of the distributions of the buyer and seller.

[1] Alternatively, a broker who tries to maximize social welfare.

We study two settings in particular: (i) the *symmetric* setting where the buyer and seller have the same valuation distribution and (ii) the *asymmetric* setting where the buyer's distribution (first-order) stochastically dominates the seller's distribution. In both settings, we show that a simple posted-pricing mechanism achieves a constant approximation to both the first-best gains-from-trade and the first-best social-welfare. Furthermore, when both the buyer's distribution and seller's distribution have monotone hazard rates, we show that this mechanism achieves a constant-approximation to the optimal profit of a posted-pricing mechanism under full information (henceforth referred to simply as the optimal profit). Interestingly, in both the symmetric and stochastic dominance settings, we observe a *strict separation* in the approximability of the first-best gains-from-trade and first-best social welfare between the setting with and without a broker.

Our positive results for the first-best gains-from-trade and first-best social-welfare stand in stark contrast to those of [148], who show pessimistic inapproximability results under a profit-maximizing broker in the full information setting. Moreover, our results are practically appealing in the sense that the broker can forgo complex data collection in governing the trade and still achieve a provable share of several optimal economic outcomes, benefiting the buyer, the seller, and the broker herself. Thus, these implications could improve the decision-making process of a governmental agency or public sector in determining whether or not to delegate the role of a mediator in a bargaining process to a third-party agent, *e.g.,* if data collection is limited to restrict the strategic broker, one can safely allow a self-interested agent to govern the trade.

In what follows, we first provide a concise version of the problem setup along with a brief summary of our technical results in Section 7.1.1. Then, in Section 7.1.2, we present a comprehensive overview of prior literature. We provide the detailed problem setup in Section 7.2, and we discuss several technical lemmas that will be used throughout our analysis in Section 7.3. Finally, we present our

results for the symmetric and the stochastic dominance settings in Sections 7.4 and 7.5, respectively. Due to the page limit, we only provide a few (rather short) proofs in the main text, and defer the rest of them to the appendix.

7.1.1 Our Results and Techniques

Overview of model. Before summarizing our results, we first briefly define our model.[2] In bilateral trade, a buyer wishes to sell a single item to a seller. The buyer's valuation v for the item is drawn from absolutely continuous distribution F, and the seller's valuation c for the item is drawn from absolutely continuous distribution G.[3] We assume that v and c are independent of each other. A broker governs the trade between the buyer and the seller, and importantly, the broker only has access to a *single* sample from each agent's distribution. Given this limited information, a natural mechanism that the broker could run is a *posted-pricing mechanism* in which she offers take-it-or-leave-it prices p and q to the buyer and seller, respectively, and the item is transferred if and only if $v \geq p$ and $c \leq q$.

Now, for any mechanism, let $x(v, c), p_b(v, c), p_s(v, c)$ denote the probability that trade occurs, the expected payment of the buyer to the broker, and the expected payment of the broker to the seller, respectively, given v and c.[4] We are interested in how the gains-from-trade (GFT), social welfare (SW), and profit (PROFIT) of the mechanism compare to the first-best gains-from-trade (FB-GFT),

[2] More details can be found in Section 7.2.

[3] By Remark 2 in [15], the assumption that F and G are absolutely continuous is without loss of generality.

[4] By the revelation principle [128], we can restrict our attention to direct Bayesian incentive-compatible mechanisms in which the buyer and the seller report their valuations truthfully.

first-best social welfare (FB-SW), and optimal profit (PROFIT*), each defined as follows:[5]

$$\text{GFT} = \mathbb{E}_{v\sim F, c\sim G}\left[(v-c)x(v,c)\right] \qquad \text{FB-GFT} = \mathbb{E}_{v\sim F, c\sim G}\left[(v-c)\mathbf{1}\{v \geq c\}\right]$$

$$\text{SW} = \mathbb{E}_{c\sim G}\left[c\right] + \text{GFT} \qquad \text{FB-SW} = \mathbb{E}_{c\sim G}\left[c\right] + \text{FB-GFT}$$

$$\text{PROFIT} = \mathbb{E}_{v\sim F, c\sim G}\left[p_b(v,c) - p_s(v,c)\right] \qquad \text{PROFIT}^* = \max_{p,q}(p-q)G(q)(1-F(p))$$

In addition, we use $\text{GFT}(p,q)$ to denote the gains-from-trade obtained from offering price p to the buyer and price q to the seller, and similarly for $\text{PROFIT}(p,q)$.

Finally, we are interested in specific problem instances characterized by the following classes of agents' distributions:

1. The buyer and seller are *symmetric* if $F = G$.
2. The buyer's distribution *(first-order) stochastically dominates* the seller's distribution if $G(x) \geq F(x)$ for all x.
3. The buyer (resp. the seller) has *monotone hazard rate* (MHR) if $(1-F(x))/F'(x)$ is nonincreasing (resp. $G(x)/G'(x)$ is nondecreasing). If the agents are symmetric, we say that $F = G$ has a *doubly monotone hazard rate* if F satisfies both of the above statements.

Overview of our results. We now present a summary of our contributions. In our paper, we consider the natural mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller,[6] and we are interested in the extent to which this mechanism approximates the first-best gains-from-trade, the first-best social welfare, and the optimal profit. Our results for the symmetric and stochastic dominance settings are summarized in Tables 7.1 and 7.2, respectively.

[5] We provide more detailed explanations for these definitions in Section 7.2.

[6] Under symmetric agents, our mechanism offers price $\max(p,q)$ to the buyer and price $\min(p,q)$ to the seller.

	With Broker			Without Broker*	
	GFT	SW	PROFIT	GFT	SW
Lower Bound	$7/24$	$2/3$	$2/55^{\dagger}$	$1/2$ [19]	$3/4$ [141]
Upper Bound	$7/24$	$2/3$	Open	$1/2$ [19]	$3/4$ [141]

*where the mechanism draws $p \sim F$ and posts p to both buyer and seller
†if F is doubly MHR

Table 7.1: Our results for the symmetric setting. Note that the approximability of the optimal profit is not applicable to the setting without a broker (or equivalently, with a benevolent broker). However, an upper bound on the approximation factor to the optimal profit remains an open problem in our setting with a broker.

We first characterize the profit of our mechanism:

Restatement of Lemma 7.3.1. *Given samples $p \sim F$ and $q \sim G$, it follows that*

$$\mathbb{E}_{p,q\sim F,G}\left[\textsc{Profit}(p,q)\right] = \frac{1}{4}\int_0^\infty G(x)^2(1-F(x))^2\,dx.$$

Next, we characterize the gains-from-trade of our mechanism, generalizing the characterization of the gains-from-trade by [141] for the single-sample regime without a broker:

Restatement of Lemma 7.3.2. *Given samples $p \sim F$ and $q \sim G$, it follows that*

$$\begin{aligned}\mathbb{E}_{p,q\sim F,G}\left[\textsc{GFT}(p,q)\right] = \frac{1}{4}\int_0^\infty G(x)^2(1-F(x))^2\,dx + \frac{1}{2}\int_0^\infty G(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q)\\ + \frac{1}{2}\int_0^\infty (1-F(p))^2\int_0^p G(x)\,dx\,dG(p).\end{aligned}$$

The proof follows from decomposing the gains-from-trade into the profit and some auxiliary terms involving integrals over $1-F(x)$ and $G(x)$, followed by some algebraic details.

Symmetric setting. With these preliminary lemmas in hand, we turn our attention to the symmetric setting. We note that the symmetric setting is well motivated in practice, *e.g.,* stock markets, treasury

markets, and dark pools. It is also largely studied in the literature, *e.g.,* [129], [19], [141], in particular when the mechanism designer only has sample access to the agents' distributions.

To begin with, we first use Lemma 7.3.2 to characterize the gains-from-trade of our single-sample mechanism in the symmetric setting:

Restatement of Lemma 7.4.1. *In the symmetric setting, given samples $p \sim F$ and $q \sim G$, our mechanism achieves*

$$\text{GFT} = \frac{1}{3}\int_0^\infty F(x)(1-F(x))\,dx - \frac{1}{6}\int_0^\infty F(x)^2(1-F(x))^2\,dx.$$

We then prove the following lower-bounds on the approximability of the first-best gains-from-trade, first-best social welfare, and optimal profit by our mechanism:

Restatement of Theorem 7.4.2, 7.4.3, and 7.4.5. *In the symmetric setting, given samples $p \sim F$ and $q \sim G$, our mechanism achieves*

$$\text{GFT} \geq \frac{7}{24}\text{FB-GFT} \text{ and } \text{SW} \geq \frac{2}{3}\text{FB-SW}.$$

Furthermore, if F is doubly MHR, then our mechanism achieves

$$\textsc{Profit} \geq \frac{2}{55}\text{FB-GFT} \geq \frac{2}{55}\textsc{Profit}^*.$$

Our proofs of the lower bounds in Theorems 7.4.2 and 7.4.3 follow rather easily from Lemma 7.4.1, although these bounds turn out to be tight (see below) despite their simplicity. On the other hand, our proof of Theorem 7.4.5 applies a result from [148] that states that when F is doubly MHR, we only

lose a constant fraction of FB-GFT $= \int_0^\infty F(x)(1-F(x))\,dx$ when we integrate over fixed quantiles of F.

Note that MHR distributions are widely adopted in single-sample mechanism design [149] and more broadly in the mechanism design literature, *e.g.,* [128], [134], [122], [123]. In particular, [14] characterize the optimal mechanism under doubly MHR distributions if the broker knows the distributions of the buyer and seller.

Now, taking a closer look at the proof of Theorem 7.4.2, we observe that our lower bound can be made arbitrarily tight by taking $F(x) \approx 1/2$. A similar observation holds for the proof of Theorem 7.4.3. Following this intuition, we obtain the following matching upper bounds on the approximability of the first-best gains-from-trade and first-best social welfare by our mechanism:

Restatement of Theorem 7.4.8 and 7.4.9. *For any $\epsilon > 0$, there exists a symmetric instance of the bilateral trade problem such that our mechanism achieves*[7]

$$\text{GFT} < \left(\frac{7}{24} + \epsilon\right) \text{FB-GFT}, \ \text{SW} < \left(\frac{2}{3} + \epsilon\right) \text{FB-SW}.$$

Asymmetric setting. In addition to the symmetric setting, we consider the setting where the buyer's distribution stochastically dominates the seller's distribution. This setting is widely adopted in the literature, *e.g.,* [150], [151], and is in fact one of the most standard assumptions for asymmetric distributions when the mechanism designer has limited information [19].

First, we use Lemma 7.3.2 to lower-bound the gains-from-trade of our single-sample mechanism:

[7] We remark that these upper bounds hold for our particular single-sample mechanism, not for arbitrary single-sample mechanisms. The upper bounds by [19] and [141] in single-sample bilateral trade without a broker hold in a similar sense, since there has been no proven information-theoretic barrier for sample-based mechanisms for bilateral trade, to the best of our knowledge.

	With Broker			Without Broker	
	GFT	SW	PROFIT	GFT*	SW†
Lower Bound	0.1254	0.1321	$1/180$‡	$1/4$ [19]	$1/2$ [30] §
Upper Bound	$7/48$	$1/6$	Open	$7/16$ [19]	Open

*where the mechanism draws $p \sim F$ and posts p to both buyer and seller
†where the mechanism draws $q \sim G$ and posts q to both buyer and seller
‡if F and G have monotone hazard rates
§This lower bound applies to arbitrary distributions, even without the stochastic dominance assumption.

Table 7.2: Our results for the setting where F stochastically dominates G. An upper bound on the approximability of the first-best social-welfare in the stochastic dominance setting without a broker is unresolved to the best of our knowledge, although [30] show an upper bound of $1/2$ for general distributions in the setting without a broker. However, since [30] use an optimization problem to obtain the exact worst-case approximation ratio, it is unclear whether there exists a problem instance where F stochastically dominates G that achieves the $1/2$ bound.

Restatement of Lemma 7.5.1. *Suppose F first-order stochastically dominates G. Then, given samples $p \sim F$ and $q \sim G$, our mechanism achieves*

$$\text{GFT} \geq \frac{1}{12} \int_0^\infty \left(3G(x)^2(1-F(x))^2 + 2F(x)^3(1-F(x)) + 2(1-G(x))^3 G(x)\right) dx.$$

Analogous to the symmetric setting, we then prove the following lower bounds on the approximability of the first-best gains-from-trade, first-best social welfare, and optimal profit achieved by our mechanism:

Restatement of Theorem 7.5.2, 7.5.3, and 7.5.4. *Suppose F first-order stochastically dominates G. Then, given samples $p \sim F$ and $q \sim G$, our mechanism achieves*

$$\text{GFT} \geq 0.1254 \cdot \text{FB-GFT}, \ \text{SW} \geq 0.1321 \cdot \text{FB-SW}.$$

Furthermore, if F and G both have monotone hazard rates, then our mechanism achieves

$$\textsc{Profit} \geq \frac{1}{180}\text{FB-GFT} \geq \frac{1}{180}\textsc{Profit}^*.$$

Our proofs of Theorems 7.5.2 and 7.5.3 use the fact that Lemma 7.5.1 reduces the problem of lower-bounding the approximation ratios to the first-best gains-from-trade and first-best social welfare of our mechanism to finding the minima of two optimization problems. On the other hand, our proof of Theorem 7.5.4 adapts a technique from [148] to the stochastic dominance setting to lower-bound the profit of our mechanism by the first-best gains-from-trade, exploiting the MHR property to obtain our result.

Similar to the symmetric setting, the proofs of Theorem 7.5.2 and 7.5.3 suggest that we can prove corresponding upper bounds by taking $F(x) \approx \alpha_F$ and $G(x) \approx \alpha_G$ for suitably chosen constants α_F and α_G. However, the analysis for the stochastic dominance setting is much more technical than that for the symmetric setting, as upper-bounding the gains-from-trade lost in the inequalities in the proof of Lemma 7.5.1 turns out to be significantly more challenging. Nevertheless, we show the following upper bounds:

Restatement of Theorem 7.5.5 and 7.5.6. *For any $\epsilon > 0$, there exists F and G, where F first-order stochastically dominates G, such that our mechanism achieves*

$$\text{GFT} < \left(\frac{7}{48} + \epsilon\right)\text{FB-GFT}, \ \text{SW} < \left(\frac{1}{6} + \epsilon\right)\text{FB-SW}.$$

We remark that in both single-sample settings, there is a strict separation between the approximability of the first-best gains-from-trade and first-best social welfare with vs. without a broker, at

least among posted-pricing mechanisms whose prices must come from the samples they draw. Notably, the relative gaps between our upper and lower bound are at least as small as the corresponding relative gaps in the setting without a broker.

7.1.2 Related Work

The bilateral trade problem was introduced in the seminal work of [14], who show that no incentive-compatible, individually-rational, and budget-balanced mechanism can achieve ex-post efficiency. As such, later work has focused on the *approximability* of the first-best efficiency by such mechanisms. In fact, much of the literature has considered the approximability of the first-best efficiency by *posted-pricing mechanisms*, due the simplicity of these mechanisms ([15, 30, 120, 122, 123, 141]).

Notably, all the literature above assumes that the trade is governed by a benevolent principal. However, in the same work where they introduce bilateral trade, [14] introduce a variation of bilateral trade where a profit-maximizing broker governs the trade. In spite of this natural extension, there had not been much work in this direction until [148] initiated the study of the approximability of the first-best gains-from-trade in bilateral trade with a broker.

Orthogonally, other works have considered markets under limited information. For example, [149] initiated the study of revenue maximization in one-sided markets where the mechanism has access to a single sample from the buyer's distribution, with follow-up works by [152], [153], and [154]. Analogously, [19], [141] and [30] study the approximability of the first-best efficiency in bilateral trade when the mechanism only has access to a single sample from the distribution of each agent. These works have essentially resolved the worst-case approximation ratios of the first-best gains-from-trade and

first-best social welfare in the symmetric and stochastic dominance settings for single-sample bilateral trade without a broker, and our paper draws an analogy to this line of work in single-sample bilateral trade *with a broker*.

7.2 Model

Bilateral trade with a broker. In bilateral trade, a buyer wishes to sell a single item to a seller. The valuations v and c of the buyer and seller, respectively, are drawn from distributions F and G, respectively, both of which are bounded absolutely-continuous distributions supported on nonnegative real values.[8] We assume that v and c are independent of each other. The buyer and seller report valuations $\tilde{v}$ and $\tilde{c}$, respectively, to the broker, who determines whether the item is transferred, and if so, how much each agent pays or receives.[9] Importantly, the broker only has access to a *single* sample from each agent's distribution on which to base her decisions. More formally, the broker runs a mechanism $\textsc{Mec} = (x, p_b, p_s)$, where given the agents' reported valuations $\tilde{v}$ and $\tilde{c}$, $x(\tilde{v}, \tilde{c}), p_b(\tilde{v}, \tilde{c})$, and $p_s(\tilde{v}, \tilde{c})$ are the probability that the item is transferred, the expected payment of the buyer to the broker, and the expected payment of the broker to the seller, respectively.

Desiderata. Among the set of all possible mechanisms, we are especially interested in mechanisms that have several nice properties. First, we say that a mechanism $\textsc{Mec} = (x, p_b, p_s)$ is *(Bayes-Nash) incentive compatible* (BNIC) if it is in each agent's best interest to report his valuation truthfully to the

[8] By Remark 2 in [15], the assumption that F and G are absolutely continuous is without loss of generality.

[9] Although an arbitrary mechanism may allow the buyer and seller to interact with the broker in other ways, since we only consider incentive-compatible mechanisms in our paper, by the revelation principle [128], we can consider direct mechanisms in which the agents directly report their valuations to the broker without any loss of generality.

broker in expectation, assuming truthful reporting by the other agent, *i.e.,*

$$\mathbb{E}_{c\sim G}\left[vx(v,c) - p_b(v,c)\right] \geq \mathbb{E}_{c\sim G}\left[vx(\tilde{v},c) - p_b(\tilde{v},c)\right]$$

and

$$\mathbb{E}_{v\sim F}\left[p_s(v,c) - cx(v,c)\right] \geq \mathbb{E}_{v\sim F}\left[p_s(v,\tilde{c}) - cx(v,\tilde{c})\right]$$

for all $v, c, \tilde{v}, \tilde{c}$. We further say that a mechanism is *dominant-strategy incentive compatible* (DSIC) if it is in each agent's best interest to report his valuation truthfully to the broker, *regardless of the other agent's bid*, *i.e.,*

$$vx(v,c) - p_b(v,c) \geq vx(\tilde{v},c) - p_b(\tilde{v},c)$$

and

$$p_s(v,c) - cx(v,c) \geq p_s(v,\tilde{c}) - cx(v,\tilde{c})$$

for all $v, c, \tilde{v}, \tilde{c}$.

Now, to incentivize the agents to participate in the mechanism, the agents should not expect to lose anything as a result of their participation. To this end, we say that MEC is *individually-rational* (IR) if

$$\mathbb{E}_{v\sim F, c\sim G}\left[vx(\tilde{v},\tilde{c}) - p_b(\tilde{v},\tilde{c})\right] \geq 0$$

and

$$\mathbb{E}_{v\sim F, c\sim G}\left[p_s(\tilde{v},\tilde{c}) - cx(\tilde{v},\tilde{c})\right] \geq 0.$$

We further say that MEC is *ex-post individually-rational* if the above inequalities hold pointwise, *i.e.,*

$$vx(\tilde{v},\tilde{c}) - p_b(\tilde{v},\tilde{c}) \geq 0$$

and

$$p_s(\tilde{v},\tilde{c}) - cx(\tilde{v},\tilde{c}) \geq 0$$

for all v, c.

In light of the above desiderata, we will focus on a specific class of mechanisms, known as *posted-pricing mechanisms*, in which the broker offers (possibly randomized) take-it-or-leave-it prices p and q to the buyer and seller, respectively, and the item is exchanged if and only if $v \geq p$ and $c \leq q$. Note that posted-pricing mechanisms are clearly DSIC[10] and ex-post IR.

Metrics and benchmarks. Throughout, we consider three benchmarks for mechanism MEC $= (x, p_b, p_s)$. First, we define the *gains-from-trade* (GFT) of a mechanism as the expected marginal increase in the utility of society from the trade, *i.e.,*

$$\text{GFT} = \mathbb{E}_{v\sim F, c\sim G}\left[x(\tilde{v},\tilde{c})(v-c)\right],$$

[10]Thus, we write v and c instead of $\tilde{v}$ and $\tilde{c}$ for posted-pricing mechanisms, since the agents have no incentive to misreport their valuations.

and we define the *social welfare* (SW) of a mechanism as the expected utility of a society from the trade, *i.e.,*

$$\text{SW} = \mathbb{E}_{c\sim C}\left[c\right] + \text{GFT}.$$

Our benchmarks for these two measures of efficiency are the gains-from-trade (resp. social welfare) obtained by an *ex-post efficient* mechanism, *i.e.,* a mechanism that always enforces trade whenever the buyer's valuation v weakly exceeds the seller's valuation c. Thus, we define the first-best gains-from-trade (FB-GFT) as

$$\text{FB-GFT} = \mathbb{E}_{v\sim F, c\sim G}\left[(v-c)\mathbf{1}\{v \geq c\}\right]$$

and the first-best social-welfare (FB-SW) as

$$\text{FB-SW} = \mathbb{E}_{c\sim C}\left[c\right] + \text{FB-GFT}.$$

Finally, we define the profit (PROFIT) of a mechanism to be

$$\text{PROFIT} = \mathbb{E}_{v\sim F, c\sim G}\left[p_b(\tilde{v},\tilde{c}) - p_b(\tilde{v},\tilde{c})\right],$$

and we will be comparing the profit of our mechanisms to the optimal profit PROFIT* of any posted-pricing mechanism under full information, *i.e.,*

$$\text{PROFIT}^* = \max_{p,q}(p-q)G(q)(1-F(p)).$$

As one final piece of notation, we use $\text{GFT}(p, q)$ to denote the gains-from-trade obtained from offering price p to the buyer and price q to the seller, and similarly for $\textsc{Profit}(p, q)$.

Distribution types. Lastly, we introduce some relevant definitions related to the agents' distributions F and G. We say that the buyer and seller are *symmetric* if $F = G$, and we say that the buyer's distribution F (first-order) *stochastically dominates* the seller's distribution G if $G(x) \geq F(x)$ for all $x \geq 0$. The buyer's distribution satisfies the *monotone hazard rate* (MHR) property if $(1 - F(x))/F'(x)$ is nonincreasing, and the seller's distribution satisfies the MHR property if $G(x)/G'(x)$ is nondecreasing. Finally, in the symmetric case where $F = G$, we say that F has a *doubly monotone hazard rate* if both of the above conditions hold for F.

7.3 Preliminary Lemmas

Hereafter, we consider the mechanism draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller, and we are interested in the extent to which this mechanism approximates the first-best gains-from-trade, the first-best social welfare, and the optimal profit. To this end, we first prove the following characterization of the profit of our mechanism:

Lemma 7.3.1. *A mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller achieves*

$$\mathbb{E}_{p,q\sim F,G}\left[\textsc{Profit}(p, q)\right] = \frac{1}{4}\int_0^\infty G(x)^2(1 - F(x))^2\,dx.$$

Proof. We have that

$$\begin{aligned}\mathbb{E}_{p,q\sim F,G}\left[\textsc{Profit}(p,q)\right] &= \int_0^\infty \int_q^\infty (p-q)(1-F(p))G(q)\,dF(p)\,dG(q)\\ &= \int_0^\infty G(q)\left(\int_q^\infty (p-q)(1-F(p))\,dF(p)\right)dG(q)\\ &= \int_0^\infty G(q)\left(\int_q^\infty (p-q)(1-F(p))F'(p)\,dp\right)dG(q).\end{aligned}$$

Integrating by parts over $(1-F(p))F'(p)$ and $(p-q)$, we can further expand by

$$\begin{aligned}\mathbb{E}_{p,q\sim F,G}\left[\textsc{Profit}(p,q)\right] &= \int_0^\infty G(q)\left(-\frac{(1-F(p))^2}{2}(p-q)\Big|_q^\infty + \frac{1}{2}\int_q^\infty (1-F(p))^2\,dp\right)dG(q)\\ &= \frac{1}{2}\int_0^\infty G(q)\left(\int_q^\infty (1-F(p))^2\,dp\right)dG(q)\\ &= \frac{1}{2}\int_0^\infty G(q)\left(\int_q^\infty (1-F(p))^2\,dp\right)G'(q)\,dq.\end{aligned}$$

Integrating by parts again over $G(q)G'(q)$ and $(1-F(p))^2$, we finally obtain

$$\begin{aligned}\mathbb{E}_{p,q\sim F,G}\left[\textsc{Profit}(p,q)\right] &= \frac{1}{2}\left(\frac{G(q)^2}{2}\int_q^\infty (1-F(p))^2\,dp\Big|_0^\infty + \frac{1}{2}\int_0^\infty G(q)^2(1-F(q))^2\,dq\right)\\ &= \frac{1}{4}\int_0^\infty G(x)^2(1-F(x))^2\,dx.\end{aligned}$$

a

Next, we prove the following characterization the gains-from-trade of our mechanism:

Lemma 7.3.2. *A mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and*

price q to the seller achieves

$$\mathbb{E}_{p,q\sim F,G}\left[\text{GFT}(p,q)\right] = \frac{1}{4}\int_0^\infty G(x)^2(1-F(x))^2\,dx + \frac{1}{2}\int_0^\infty G(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q) + \frac{1}{2}\int_0^\infty (1-F(p))^2\int_0^p G(x)\,dx\,dG(p).$$

As the proof of Lemma 7.3.2 is rather algebra-heavy, we defer it to Appendix F.1.1.

7.4 Symmetric Case

In the symmetric case, there is a simple modification we can make on our mechanism to boost the gains-from-trade, social welfare, and profit. In particular, given samples $p, q \sim F$, our mechanism will offer price $\max(p,q)$ to the buyer and price $\min(p,q)$ to the seller, instead of simply offering price p to the buyer and price q to the seller. We characterize the gains-from-trade of this mechanism in the following lemma:

Lemma 7.4.1. *In the symmetric setting, a mechanism that draws samples $p, q \sim F$ and offers price $\max(p,q)$ to the buyer and price $\min(p,q)$ to the seller achieves*

$$\text{GFT} = \frac{1}{3}\int_0^\infty F(x)(1-F(x))\,dx - \frac{1}{6}\int_0^\infty F(x)^2(1-F(x))^2\,dx.$$

Proof of Lemma 7.4.1. We have that

$$\begin{aligned}
\mathrm{GFT} &= \mathbb{E}_{p,q\sim F}\left[\mathrm{GFT}(p,q)\mathbf{1}\{p \geq q\}\right] + \mathbb{E}_{p,q\sim F}\left[\mathrm{GFT}(q,p)\mathbf{1}\{q \geq p\}\right] \\
&= 2\mathbb{E}_{p,q\sim F}\left[\mathrm{GFT}(p,q)\mathbf{1}\{p \geq q\}\right] \\
&= 2\mathbb{E}_{p,q\sim F}\left[\mathrm{GFT}(p,q)\right] \\
&= \frac{1}{2}\int_0^\infty F(x)^2(1-F(x))^2\,dx + \int_0^\infty F(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q) \\
&+ \int_0^\infty (1-F(p))^2\int_0^p F(x)\,dx\,dF(p),
\end{aligned}$$

where the last equality holds by Lemma 7.3.2. Evaluating the first integral, we obtain

$$\begin{aligned}
\int_0^\infty F(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q) &= \int_0^\infty F(q)^2\int_q^\infty (1-F(x))\,dx\,F'(q)\,dq \\
&= \frac{F(q)^3}{3}\int_q^\infty (1-F(x))\,dx\Big|_0^\infty + \frac{1}{3}\int_0^\infty F(q)^3(1-F(q))\,dq \\
&\quad \text{(integrating by parts over } F(q)^2F'(q) \text{ and } \int_q^\infty (1-F(x))\,dx) \\
&= \frac{1}{3}\int_0^\infty F(x)^3(1-F(x))\,dx,
\end{aligned}$$

and evaluating the second integral, we obtain

$$\begin{aligned}
\int_0^\infty (1-F(p))^2\int_0^p F(x)\,dx\,dF(p) &= \int_0^\infty (1-F(p))^2\int_0^p F(x)\,dx\,F'(p)\,dp \\
&= -\frac{(1-F(p))^3}{3}\int_0^p F(x)\,dx\Big|_0^\infty + \frac{1}{3}\int_0^\infty (1-F(p))^3F(p)\,dp \\
&\quad \text{(integrating by parts over } (1-F(p))^2F'(p) \text{ and } \int_0^p F(x)\,dx) \\
&= \frac{1}{3}\int_0^\infty (1-F(x))^3F(x)\,dx.
\end{aligned}$$

Thus, we have that

$$\begin{aligned}\text{GFT} &= \frac{1}{2}\int_0^\infty F(x)^2(1-F(x))^2\,dx + \frac{1}{3}\int_0^\infty F(x)^3(1-F(x))\,dx + \frac{1}{3}\int_0^\infty (1-F(x))^3F(x)\,dx \\ &= \frac{1}{3}\int_0^\infty F(x)(1-F(x))\,dx - \frac{1}{6}\int_0^\infty F(x)^2(1-F(x))^2\,dx.\end{aligned}$$

a

7.4.1 Lower Bounds

The characterization in Lemma 7.3.2 leads naturally to the following simple lower bound on the approximation ratio of our mechanism to the first-best gains-from-trade:

Theorem 7.4.2. *In the symmetric setting, a mechanism that draws samples $p, q \sim F$ and offers price $\max(p,q)$ to the buyer and price $\min(p,q)$ to the seller achieves*

$$\text{GFT} \geq \frac{7}{24}\text{FB-GFT}.$$

Proof. By Lemma 7.4.1, we have that

$$\begin{aligned}\text{GFT} &= \frac{1}{3}\int_0^\infty F(x)(1-F(x))\,dx - \frac{1}{6}\int_0^\infty F(x)^2(1-F(x))^2\,dx \\ &\geq \frac{1}{3}\int_0^\infty F(x)(1-F(x))\,dx - \frac{1}{6}\left(\frac{1}{4}\int_0^\infty F(x)(1-F(x))\,dx\right) \\ &= \frac{7}{24}\int_0^\infty F(x)(1-F(x))\,dx \\ &= \frac{7}{24}\text{FB-GFT},\end{aligned}$$

where the first inequality holds since $F(x)(1-F(x)) \le 1/4$ for all x. a

Next, by using the fact that $\text{SW} = \mathbb{E}_{c\sim F}[c] + \text{GFT}$, we can apply Lemma 7.3.2 followed by some algebraic manipulations to obtain the following lower bound on the approximation ratio of our mechanism to the first-best social welfare, whose proof we defer to Appendix F.3.1:

Theorem 7.4.3. *In the symmetric setting, a mechanism that draws samples $p, q \sim F$ and offers price $\max(p,q)$ to the buyer and price $\min(p,q)$ to the seller achieves*

$$\text{SW} \ge \frac{2}{3}\text{FB-SW}.$$

Finally, we prove the following lower bound on the approximation ratio of our mechanism to the optimal profit when F is doubly MHR. We require the following result from the proof of Theorem 1.3 of [148]:

Lemma 7.4.4. *[148] Suppose F is doubly MHR. Then, if $0 < \beta \le \alpha < 1, C \ge 1$ satisfy*

$$\frac{C-1}{2}\min\left(\frac{\alpha^2-\beta^2}{2}, \alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right) \ge \max\left(\beta-\frac{\beta^2}{2}, \frac{1-\alpha^2}{2}\right),$$

we have that

$$\int_{F^{-1}(\beta)}^{F^{-1}(\alpha)} F(x)(1-F(x))\,dx \ge \frac{1}{C}\text{FB-GFT}.$$

Theorem 7.4.5. *Suppose F is doubly MHR. Then, a mechanism that draws samples $p, q \sim F$ and*

offers price $\max(p, q)$ *to the buyer and price* $\min(p, q)$ *to the seller achieves*

$$\textsc{Profit} \geq \frac{2}{55}\text{FB-GFT} \geq \frac{2}{55}\textsc{Profit}^*.$$

Proof. Suppose we have constants $0 < \beta \leq \alpha < 1, C \geq 1$ such that

$$\frac{C-1}{2}\min\left(\frac{\alpha^2-\beta^2}{2}, \alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right) \geq \max\left(\beta-\frac{\beta^2}{2}, \frac{1-\alpha^2}{2}\right).$$

Then, since F is doubly-MHR, by Lemma 7.4.4, we have that

$$\int_{F^{-1}(\beta)}^{F^{-1}(\alpha)} F(x)(1-F(x))\,dx \geq \frac{1}{C}\text{FB-GFT},$$

so

$$\begin{aligned}
\textsc{Profit} &= \mathbb{E}_{p,q\sim F}\left[\textsc{Profit}(p,q)\mathbf{1}\{p \geq q\}\right] + \mathbb{E}_{p,q\sim F}\left[\textsc{Profit}(q,p)\mathbf{1}\{q \geq p\}\right] \\
&= 2\mathbb{E}_{p,q\sim F}\left[\textsc{Profit}(p,q)\mathbf{1}\{p \geq q\}\right] \\
&= 2\mathbb{E}_{p,q\sim F}\left[\textsc{Profit}(p,q)\right] \\
&\geq \frac{1}{2}\int_{F^{-1}(\beta)}^{F^{-1}(\alpha)} F(x)^2(1-F(x))^2\,dx \\
&\geq \frac{\min(\alpha(1-\alpha), \beta(1-\beta))}{2}\int_{F^{-1}(\beta)}^{F^{-1}(\alpha)} F(x)(1-F(x))\,dx \\
&\geq \frac{\min(\alpha(1-\alpha), \beta(1-\beta))}{2C}\text{FB-GFT},
\end{aligned}$$

where the first inequality holds by Lemma 7.3.1. By (roughly) numerically optimizing the parameters

α, β, and C, we take $\alpha = 4/5, \beta = 1/5$, and $C = 11/5$ to obtain that

$$\textsc{Profit} \geq \frac{2}{55}\text{FB-GFT} \geq \frac{2}{55}\textsc{Profit}^*.$$

a

7.4.2 Upper Bounds

Observe that in the proof of the lower bound in Theorem 7.4.2, the only inequality holds with equality if and only if $F(x)(1 - F(x)) = 1/4$, or equivalently $F(x) = 1/2$, almost everywhere in the support of F. This suggests that we can prove a matching upper bound by taking $F(x) \approx 1/2$.

To construct a distribution H such that $H(x) \approx \alpha$ for $0 < \alpha < 1$, we first require the following definition:

Definition 7.4.6. Let $0 < \alpha < 1$ and $0 < \delta < \dfrac{\alpha}{2\alpha + 1}$. Define $H_{\alpha,\delta,\text{QUAD}} : [0, \delta] \to \mathbb{R}$ as

$$H_{\alpha,\delta,\text{QUAD}}(x) = \frac{\alpha - \delta}{\delta}x + \frac{(2\alpha + 1)\delta - \alpha}{\delta^2(1 - 2\delta)}x(x - \delta)$$

and $H_{\alpha,\delta,\text{LIN}} : [\delta, 1 - \delta] \to \mathbb{R}$ as

$$H_{\alpha,\delta,\text{LIN}}(x) = \alpha + \frac{2\delta}{1 - 2\delta}\left(x - \frac{1}{2}\right).$$

We now define our desired class of distributions:

Definition 7.4.7. Let $0 < \alpha < 1$ and $0 < \delta < \min\left(\dfrac{\alpha}{2\alpha + 1}, \dfrac{1 - \alpha}{3 - 2\alpha}\right)$. We define $H_{\alpha,\delta} : [0, 1] \to \mathbb{R}$

as

$$H_{\alpha,\delta}(x) = \begin{cases} H_{\alpha,\delta,\text{QUAD}}(x) & \text{if } 0 \le x \le \delta \\ H_{\alpha,\delta,\text{LIN}}(x) & \text{if } \delta \le x \le 1-\delta \\ 1 - H_{1-\alpha,\delta,\text{QUAD}}(1-x) & \text{if } 1-\delta \le x \le 1. \end{cases}$$

We prove that $H_{\alpha,\delta}$ is a differentiable CDF in Appendix F.2.

Using the above class of distributions, we prove the following upper bound on the approximation ratio to the first-best gains-from-trade of our mechanism:

Theorem 7.4.8. *For any $\epsilon > 0$, there exists a symmetric instance of the bilateral trade problem such that a mechanism that draws samples $p, q \sim F$ and offers price $\max(p, q)$ to the buyer and price $\min(p, q)$ to the seller achieves*

$$\text{GFT} < \left(\frac{7}{24} + \epsilon\right) \text{FB-GFT}.$$

Proof of Theorem 7.4.8. Let $F = H_{1/2,\delta}$ for $0 < \delta < 1/4$. Now, note that $F(x) = 1 - F(1-x)$ for

all $0 \leq x \leq 1/2$, so by symmetry, we have that

$$\begin{aligned}
&\int_0^1 F(x)(1-F(x))\left(\frac{1}{4} - F(x)(1-F(x))\right) dx \\
&= 2\int_0^{1/2} F(x)(1-F(x))\left(\frac{1}{4} - F(x)(1-F(x))\right) dx \\
&\leq \frac{1}{2}\int_0^{1/2} \left(\frac{1}{4} - F(x)(1-F(x))\right) dx \\
&\leq \frac{1}{2}\left(\int_0^{\delta} \left(\frac{1}{4} - F(x)(1-F(x))\right) dx + \int_{\delta}^{1/2} \left(\frac{1}{4} - F(x)(1-F(x))\right) dx\right) \\
&= \frac{1}{2}\left(\frac{1}{4}\delta + \int_{\delta}^{1/2} \left(\frac{1}{4} - F(x)(1-F(x))\right) dx\right).
\end{aligned}$$

Using the fact that $1/2 - \delta \leq F(x) \leq 1/2$ on $[0, 1/2]$, we upper-bound the above integral to obtain

$$\begin{aligned}
\int_{\delta}^{1/2} \left(\frac{1}{4} - F(x)(1-F(x))\right) dx &\leq \int_{\delta}^{1/2} \left(\frac{1}{4} - \frac{1}{2}\left(\frac{1}{2} - \delta\right)\right) dx \\
&= \frac{1}{2}\delta\left(\frac{1}{2} - \delta\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
\int_0^1 F(x)(1-F(x))\left(\frac{1}{4} - F(x)(1-F(x))\right) dx &\leq \frac{1}{2}\left(\frac{1}{4}\delta + \frac{1}{2}\delta\left(\frac{1}{2} - \delta\right)\right) \\
&\xrightarrow{\delta \to 0} 0,
\end{aligned}$$

so

$$\int_0^1 F(x)(1-F(x))\left(\frac{1}{4} - F(x)(1-F(x))\right) dx \xrightarrow{\delta \to 0} 0. \tag{7.1}$$

In addition,

$$\begin{aligned}
\text{FB-GFT} &\geq \mathbb{E}_{v,c\sim F}\left[(v-c)\mathbf{1}\left\{v \geq 1-\delta, c \leq \delta\right\}\right] \\
&\geq \mathbb{E}_{v,c\sim F}\left[(1-2\delta)\mathbf{1}\left\{v \geq 1-\delta, c \leq \delta\right\}\right] \\
&\geq (1-2\delta)\mathbb{P}_{v\sim F}\left[v \geq 1-\delta\right]\mathbb{P}_{c\sim G}\left[c \leq \delta\right] \\
&\geq (1-2\delta)\left(\frac{1}{2}-\delta\right)^2 \\
&\geq \frac{1}{2}\cdot\left(\frac{1}{4}\right)^2, \qquad (7.2)
\end{aligned}$$

where the last inequality holds since $\delta < 1/4$. In particular, FB-GFT is bounded away from 0. Finally, by Lemma 7.4.1,

$$\begin{aligned}
\text{GFT} &= \frac{1}{3}\int_0^1 F(x)(1-F(x))\,dx - \frac{1}{6}\int_0^1 F(x)^2(1-F(x))^2\,dx \\
&= \frac{1}{3}\int_0^1 F(x)(1-F(x))\,dx - \frac{1}{6}\left(\frac{1}{4}\int_0^1 F(x)(1-F(x))\,dx\right) \\
&\quad + \frac{1}{6}\int_0^1 F(x)(1-F(x))\left(\frac{1}{4} - F(x)(1-F(x))\right)dx. \qquad (7.3)
\end{aligned}$$

Hence, by Equations (7.1), (7.2), and (7.3),

$$\frac{\text{GFT}}{\text{FB-GFT}} \xrightarrow{\delta\to 0} \frac{7}{24}.$$

a

Similarly, by observing the proof of the lower bound in Theorem 7.4.3, we construct a (simpler) problem instance to show the following matching upper bound:

Theorem 7.4.9. *For any $\epsilon > 0$, there exists a symmetric instance of the bilateral trade problem such that a mechanism that draws samples $p, q \sim F$ and offers price $\max(p, q)$ to the buyer and price $\min(p, q)$ to the seller achieves*

$$\text{SW} < \left(\frac{2}{3} + \epsilon\right) \text{FB-SW}.$$

Proof. Let $F(x) = x^\delta$ for $\delta > 0$. Then,

$$\begin{aligned}
\text{FB-SW} &= \mathbb{E}_{c\sim F}[c] + \text{FB-GFT} \\
&= \int_0^1 (1 - F(x))\,dx + \int_0^1 F(x)(1 - F(x))\,dx \\
&= \int_0^1 (1 - F(x)^2)\,dx \\
&= \int_0^1 (1 - x^{2\delta})\,dx \\
&= 1 - \frac{1}{2\delta + 1} \\
&= \frac{2\delta}{2\delta + 1}.
\end{aligned}$$

In addition, as in the proof of Theorem 7.4.3, we have that

$$\begin{aligned}
\text{SW} &= \frac{1}{6}\left(\int_0^1 (1 - F(x)^2)(F(x)^2 - 2F(x) + 4)\,dx + 2\int_0^1 (1 - F(x))\,dx\right) \\
&= \frac{1}{6}\left(\int_0^1 (4 - 2x^\delta - 3x^{2\delta} + 2x^{3\delta} - x^{4\delta})\,dx + 2\int_0^1 (1 - x^\delta)\,dx\right) \\
&= \frac{1}{6}\left(\left(4 - \frac{2}{\delta + 1} - \frac{3}{2\delta + 1} + \frac{2}{3\delta + 1} - \frac{1}{4\delta + 1}\right) + 2\left(1 - \frac{1}{\delta + 1}\right)\right) \\
&= 1 - \frac{2/3}{\delta + 1} - \frac{1/2}{2\delta + 1} + \frac{1/3}{3\delta + 1} - \frac{1/6}{4\delta + 1}.
\end{aligned}$$

Now, letting $h(\delta) = (\delta+1)(2\delta+1)(3\delta+1)(4\delta+1)$, we obtain[11]

$$\begin{aligned}\text{SW} &= \frac{1+10\delta+O(\delta^2)}{h(\delta)} - \frac{2(1+9\delta+O(\delta^2))}{3h(\delta)} - \frac{1+8\delta+O(\delta^2)}{2h(\delta)} \\ &\quad + \frac{1+7\delta+O(\delta^2)}{3h(\delta)} - \frac{1+6\delta+O(\delta^2)}{6h(\delta)} \\ &= \frac{4\delta+O(\delta^2)}{3h(\delta)}.\end{aligned}$$

Thus,

$$\frac{\text{SW}}{\text{FB-SW}} = \frac{2+O(\delta)}{3(\delta+1)(3\delta+1)(4\delta+1)} \xrightarrow{\delta\to 0} \frac{2}{3}.$$

a

7.5 Stochastic Dominance Case

Recall that we are interested in the extent to which the mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller approximates the first-best gains-from-trade, the first-best social welfare, and the optimal profit. We now consider the setting in which the buyer's distribution F stochastically dominates the seller's distribution G.

7.5.1 Lower Bounds

First, we prove the following lower bound on the gains-from-trade of our mechanism (see Appendix F.1.2):

[11] Throughout, we use big-O notation to hide dominated terms with respect to δ.

Lemma 7.5.1. *Suppose F first-order stochastically dominates G. Then, a mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller achieves*

$$\text{GFT} \geq \frac{1}{12} \int_0^\infty (3G(x)^2(1-F(x))^2 + 2F(x)^3(1-F(x)) + 2(1-G(x))^3 G(x))\, dx.$$

Next, by lower-bounding the ratio between the integrand in Lemma 7.5.1 and the integrand of the first-best gains-from-trade (*i.e.,* $G(x)(1-F(x))$), we obtain the following lower bound on the approximation ratio of our mechanism to the first-best gains-from-trade:

Theorem 7.5.2. *Suppose F first-order stochastically dominates G. Then, a mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller achieves*

$$\text{GFT} \geq C \cdot \text{FB-GFT},$$

where $C \approx 0.1254$.

Proof. Consider the following optimization problem:

$$\begin{aligned} \underset{f,g}{\text{minimize}} \quad & \frac{3g^2(1-f)^2 + 2f^3(1-f) + 2(1-g)^3 g}{12g(1-f)} \\ \text{subject to} \quad & 0 \leq f \leq g \leq 1, f < 1, g > 0. \end{aligned}$$

Solving this optimization problem numerically[12], we obtain that the optimum $C \approx 0.1254$ occurs at $(f, g) \approx (0.3909, 0.6091)$. Thus,

$$\frac{1}{12}(3G(x)^2(1-F(x))^2 + 2F(x)^3(1-F(x)) + 2(1-G(x))^3 G(x)) \geq C \cdot G(x)(1-F(x))$$

[12]This is solved via WolframAlpha.

for all x, so by Lemma 7.5.1,

$$\text{GFT} \geq C \int_0^\infty G(x)(1 - F(x))\,dx = C \cdot \text{FB-GFT}.$$

a

We use an analogous technique to prove the following lower bound on the approximation ratio of our mechanism to the first-best social welfare:

Theorem 7.5.3. *Suppose F first-order stochastically dominates G. Then, a mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller achieves*

$$\text{SW} \geq \frac{3 - \sqrt{2}}{12}\text{FB-SW},$$

where $(3 - \sqrt{2})/12 \approx 0.1321$.

Proof. By Lemma 7.5.1,

$$\begin{aligned}\text{SW} &= \mathbb{E}_{c\sim G}\left[c\right] + \text{GFT} \\ &\geq \int_0^\infty (1 - G(x))\,dx + \frac{1}{12}\int_0^\infty (3G(x)^2(1 - F(x))^2 + 2F(x)^3(1 - F(x)) + 2(1 - G(x))^3G(x))\,dx.\end{aligned}$$

Now, consider the following optimization problem:

$$\begin{aligned}\underset{f,g}{\text{minimize}} \quad & \frac{12(1-g) + 3g^2(1-f)^2 + 2f^3(1-f) + 2(1-g)^3g}{12(1-fg)} \\ \text{subject to} \quad & 0 \leq f \leq g \leq 1, fg < 1.\end{aligned}$$

Solving this optimization problem numerically,[12] we obtain that the optimum $\alpha = (3-\sqrt{2})/12$ occurs

at $(f, g) = (1/\sqrt{2}, 1)$. Thus,

$$(1 - G(x)) + \frac{1}{12}(3G(x)^2(1 - F(x))^2 + 2F(x)^3(1 - F(x)) + 2(1 - G(x))^3G(x))$$
$$\geq \frac{3 - \sqrt{2}}{12}((1 - G(x)) + G(x)(1 - F(x)))$$

for all x, so

$$\text{SW} \geq \frac{3 - \sqrt{2}}{12}\left(\int_0^\infty (1 - G(x))\,dx + \int_0^\infty G(x)(1 - F(x))\,dx\right) = \frac{3 - \sqrt{2}}{12}\text{FB-SW}.$$

a

Finally, we prove a lower bound on the approximation ratio of our mechanism to the optimal profit by adapting Lemma 7.4.4 from [148] for the stochastic domination setting:

Theorem 7.5.4. *Suppose F and G are MHR distributions such that F first-order stochastically dominates G. Then, a mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller achieves*

$$\textsc{Profit} \geq \frac{1}{180}\text{FB-GFT} \geq \frac{1}{180}\textsc{Profit}^*.$$

Proof. Suppose there exist constants $0 < \beta \leq \alpha < 1, C \geq 1$ such that

$$\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)(1 - F(x))\,dx \geq \frac{1}{C}\text{FB-GFT}.$$

Then, by Lemma 7.3.1,

$$\begin{aligned}\textsc{Profit} &= \mathbb{E}_{p,q\sim F,G}\left[\textsc{Profit}(p,q)\right] \\ &\geq \frac{1}{4}\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)^2(1-F(x))^2\,dx \\ &\geq \frac{\beta(1-\alpha)}{4}\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)(1-F(x))\,dx \\ &\geq \frac{\beta(1-\alpha)}{4C}\text{FB-GFT}.\end{aligned}$$

(Note that $F^{-1}(\alpha) \geq F^{-1}(\beta) \geq G^{-1}(\beta)$, since F stochastically dominates G.)

Now, it suffices to find constants α, β, C such that

$$(C-1)\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)(1-F(x))\,dx \geq \int_0^{G^{-1}(\beta)} G(x)(1-F(x))\,dx + \int_{F^{-1}(\alpha)}^{\infty} G(x)(1-F(x))\,dx, \tag{7.4}$$

since

$$\begin{aligned}\text{FB-GFT} &= \int_0^{\infty} G(x)(1-F(x))\,dx \\ &= \int_0^{G^{-1}(\beta)} G(x)(1-F(x))\,dx + \int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)(1-F(x))\,dx + \int_{F^{-1}(\alpha)}^{\infty} G(x)(1-F(x))\,dx.\end{aligned}$$

For the integral on the left-hand side of (7.4), we have that

$$\begin{aligned}\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)(1-F(x))\,dx &= \int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)\frac{1-F(x)}{F'(x)}F'(x)\,dx \\ &\geq \frac{1-\alpha}{F'(F^{-1}(\alpha))}\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)\,dF(x) \\ &\geq \frac{1-\alpha}{F'(F^{-1}(\alpha))}\int_{F^{-1}(\beta)}^{F^{-1}(\alpha)} F(x)\,dF(x) \\ &= \frac{1-\alpha}{F'(F^{-1}(\alpha))}\left(\frac{\alpha^2-\beta^2}{2}\right),\end{aligned}$$

where the first inequality holds since F is MHR and the second inequality holds since F stochastically dominates G. At the same time, we have that

$$\begin{aligned}\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)(1-F(x))\,dx &\geq \int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} \frac{G(x)}{G'(x)}(1-F(x))G'(x)\,dx \\ &\geq \frac{\beta}{G'(G^{-1}(\beta))}\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} (1-F(x))\,dG(x) \\ &\geq \frac{\beta}{G'(G^{-1}(\beta))}\int_{G^{-1}(\beta)}^{G^{-1}(\alpha)} (1-G(x))\,dG(x) \\ &\geq \frac{\beta}{G'(G^{-1}(\beta))}\left(\alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right).\end{aligned}$$

Thus, we can lower-bound the left-hand side of (7.4) as

$$\begin{aligned}\int_{G^{-1}(\beta)}^{F^{-1}(\alpha)} G(x)(1-F(x))\,dx &\geq \frac{1}{2}\left(\frac{1-\alpha}{F'(F^{-1}(\alpha))}\left(\frac{\alpha^2-\beta^2}{2}\right)+\frac{\beta}{G'(G^{-1}(\beta))}\left(\alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right)\right) \\ &= \frac{1}{2}\min\left(\frac{\alpha^2-\beta^2}{2}, \alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right)\left(\frac{1-\alpha}{F'(F^{-1}(\alpha))}+\frac{\beta}{G'(G^{-1}(\beta))}\right).\end{aligned} \tag{7.5}$$

Now, consider the right-hand side of (7.4). Upper-bounding the first integral, we obtain

$$\begin{aligned}\int_0^{G^{-1}(\beta)} G(x)(1-F(x))\,dx &= \int_0^{G^{-1}(\beta)} \frac{G(x)}{G'(x)}(1-F(x))G'(x)\,dx \\ &\le \frac{\beta}{G'(G^{-1}(\beta))}\int_0^{G^{-1}(\beta)} (1-F(x))\,dG(x) \\ &\le \frac{\beta}{G'(G^{-1}(\beta))}\cdot\beta,\end{aligned}$$

and for the second integral, we similarly obtain

$$\begin{aligned}\int_{F^{-1}(\alpha)}^{\infty} G(x)(1-F(x))\,dx &= \int_{F^{-1}(\alpha)}^{\infty} G(x)\frac{1-F(x)}{F'(x)}F'(x)\,dx \\ &\le \frac{1-\alpha}{F'(F^{-1}(\alpha))}\int_{F^{-1}(\alpha)}^{\infty} G(x)\,dF(x) \\ &\le \frac{1-\alpha}{F'(F^{-1}(\alpha))}\cdot(1-\alpha).\end{aligned}$$

Hence, we can upper-bound the right-hand side in (7.4) as

$$\int_0^{G^{-1}(\beta)} G(x)(1-F(x))\,dx + \int_{F^{-1}(\alpha)}^{\infty} G(x)(1-F(x))\,dx \tag{7.6}$$

$$\le \max(\beta, 1-\alpha)\left(\frac{1-\alpha}{F'(F^{-1}(\alpha))} + \frac{\beta}{G'(G^{-1}(\beta))}\right). \tag{7.7}$$

Therefore, from (7.4), we see that it suffices by (7.5) and (7.7) to find constants α, β, C such that

$$\frac{C-1}{2}\min\left(\frac{\alpha^2-\beta^2}{2}, \alpha-\beta-\frac{\alpha^2-\beta^2}{2}\right) \ge \max(\beta, 1-\alpha).$$

By (roughly) numerically optimizing the parameters α, β and C here, taking $\alpha = 2/3, \beta = 1/3$, and

$C = 5$, we obtain

$$\textsc{Profit} \geq \frac{\beta(1-\alpha)}{4C}\text{FB-GFT} = \frac{1}{180}\text{FB-GFT} \geq \frac{1}{180}\textsc{Profit}^*.$$

a

7.5.2 Upper Bounds

Similar to the symmetric case, our proof of the lower bound in Theorem 7.5.2 suggests that we can prove a corresponding upper bound by taking $F(x) \approx \alpha_F$ and $G(x) \approx \alpha_G$ for α_F, α_G corresponding to the solution of the optimization problem in the proof of Theorem 7.5.2. Unfortunately, since the lower bound in Theorem 7.5.2 depends on the lower bound in Lemma 7.5.1, which might not necessarily be tight, choosing these values of α_F, α_G might not give the best bound. Thus, we instead consider $F = H_{\alpha_F,\delta}$ and $G = H_{\alpha_G,\delta}$ for $0 < \alpha_F \leq \alpha_G < 1$ *to be determined* and optimize over α_F, α_G to obtain the following upper bound on the approximation ratio of our mechanism to the first-best gains-from-trade (see Appendices F.2 and F.4):[13]

Theorem 7.5.5. *For any $\epsilon > 0$, there exists F and G, where F first-order stochastically dominates G, such that a mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller achieves*

$$\text{GFT} < \left(\frac{7}{48} + \epsilon\right)\text{FB-GFT}.$$

[13]Note that $7/48$ is exactly twice the corresponding upper bound in Theorem 7.4.8 in the symmetric case. This is not a coincidence: The problem instance that we use to prove these upper bounds turns out to be the same for both theorems. However, since our mechanism in the stochastic dominance setting posts prices p and q instead of prices $\max(p, q)$ and $\min(p, q)$, we lose half of the gains-from-trade.

Likewise, we obtain the following upper bound on the approximation ratio of our mechanism to the first-best social welfare:

Theorem 7.5.6. *For any $\epsilon > 0$, there exists F and G, where F first-order stochastically dominates G, such that a mechanism that draws samples $p \sim F, q \sim G$ and offers price p to the buyer and price q to the seller achieves*

$$\text{SW} < \left(\frac{1}{6} + \epsilon\right) \text{FB-SW}.$$

Part IV

Delegation of Information Acquisition

Large Language Models (LLMs), in the past few years, received significant attention from extensive areas with its ability to serve numerous tasks including question answering, image generation, code completion and reasoning [155–158]. Recent success of commercial LLMs such as ChatGPT, Gemini, and Claude [159, 160] gives evidence on the feasibility of LLMs as personal assistants, supporting individuals' daily life in decision making, problem solving, and planning.

To serve as a genuine personalized assistant, however, it is crucial to align the LLM's outcome with the designated human preference, which remains as a core challenge in LLMs [161–163]. Recent works suggest that aligning an LLM's output with human preference can be partially addressed by fine-tuning the model with reinforcement learning from human feedback (RLHF) [164–166] (see [167] for comprehensive exposition). This line of research, however, typically requires significant computational resources due to fine-tuning, extensive human labeling, and frequent model updates, rendering it less scalable in practical and rapidly changing environments or as a genuine personalized assistant tailored to each individual's preference.

Consequently, a more recent thread of studies proposes *decoding-time alignment* methods, aiming to align the LLM's output on the fly during inference time without fine-tuning the LLM [168–170]. While these methods hold promise in terms of scalability, their theoretical underpinnings remain largely uncharted, despite some recent evidences [171–173], and how it could sample-efficiently learn the preference from user feedback is yet questionable. Moreover, even the decoding algorithm itself lacks foundational understanding from theoretical perspective despite being central to the LLM's quality, even the simplest ones such as greedy decoding or beam search, beyond a few recent

works [174, 175].

Chapter 8: Tokenized Bandit for LLM Decoding and Alignment

This chapter presents [25] (ICML'26), which studies algorithms to effectively delegate information acquisition process to LLMs by smartly adjusting the decoding algorithm of the LLMs.

8.1 Introduction

In this work, we propose a fresh perspective in the decoding-time alignment as well as decoding algorithm for LLMs by framing the LLM alignment and decoding problem as variants of linear/multi-armed bandits framework whose utility is represented via a sequence function and when a decision maker faces an online problem to select the sequence irrevocably in a token-by-token manner. Overall, our contributions are three-folds:

- We present tokenized variants of the linear and multi-armed bandit problems, which turns out to be closely connected to bandit tree search (BTS) problem [176]. We first show that the general problem setup fundamentally exhibits exponential dependency in the maximal length of token sequences for any algorithm.

- To address it, we provide a fairly reasonable assumption, dubbed *diminishing distance with more commons* (DDMC), and provide sublinear regret algorithms for both problems. This complements the submodularity, widely adopted in sequence function maximization.

- We explain how our algorithms can be applied in the decoding-time LLM alignment. We discuss further interesting implications in the static counterpart of our problem, which might be of independent interest, demonstrating the provably optimality of the greedy decoding under DDMC structure, justifying empirical evidences [177] on the effectiveness of greedy decoding in certain tasks. We further provide an empirical evidence on DDMC for real-world datasets and justify its feasibility, though our primary focus is on theoretical implications.

In what follows, we briefly introduce our contributions and implications therein. Section 8.2 formally introduces our general problem setup. We provide our results for tokenized linear bandit problem in Section 8.3. Section 8.4 presents our results for the tokenized multi-armed bandit problem. Implications and applications to decoding-time alignment, decoding, and experimental results are provided in Section 8.5. Related works, and all the proofs are deferred to Appendix G.1 due to page limits. Appendix G.2 presents our results with weakened version of DDMC, combined with lookahead decoding, and Appendix G.3 discusses how our problems are connected to the seminal bandit tree search (BTS) problem

8.1.1 Overview of results, assumptions, and implications

We start with the most general problem setup that will subsume *tokenized linear bandit* (TLB) and *tokenized multi-armed bandit problem* (TMAB).[1] Each round $t \in [T]$, a user arrives and provides a *query* x_t as a context, and a decision maker (DM) selects token sequence $\mathbf{y}_t \in \mathcal{V}^*$[2] one at a time from a fixed vocabulary $\mathcal{V}$ with $|\mathcal{V}| = n$. Once the token sequence is finalized, the DM receives a *random utility* $r_t(\mathbf{y}_t)$ from the user, capturing how well the produced text aligns with the user's preference. The

[1] We note that TMAB does not directly reduce to TLB as a special case unlike the standard bandit problem, due to an exponential blow up in the dimension of the latent parameter. Thus, it requires independent algorithms and analysis.

[2] A^* denotes a set of finite sequence over A.

random utility consists of two parts: (i) sequence function $u_t(x_t, \mathbf{y}_t)$ that maps a pair of context and token sequence to nonnegative real value, and (ii) random noise η_t analogous to the standard bandit problems. The benchmark is to obtain sublinear regret against the optimal sequence in hindsight per each round, *i.e.,* $\textsc{Opt}_t = \mathbb{E}\left[\max_{\mathbf{y}\in\mathcal{V}^*} r_t(\mathbf{y})\right]$ and $\textsc{Opt} = \sum_{t=1}^{T} \textsc{Opt}_t$. We assume that the output token sequence is restricted to be within length of L, and seeks for algorithms with regret linear on L and sublinear on T.

In TLB, our utility function is linearly parameterized by $u_t(x_t, \mathbf{y}) = \langle \theta, e(x_t, \mathbf{y})\rangle$ given an embedding function $e : \mathcal{V}^* \times \mathcal{V}^* \to \mathbb{R}^d$ that maps a query and a token sequence to a feature vector, which can be viewed as an embedding vector that a typical LLM returns. In TMAB, utility function could be arbitrary, but the context remains the same $x_t = x$ to pose the problem learnable. For both problems, DDMC assumption is defined as follows: for any two token sequences $\mathbf{y}, \mathbf{z} \in \mathcal{V}^*$ with the same length $|\mathbf{y}| = |\mathbf{z}|$ and a token $\tau \in \mathcal{V}$, it follows that

$$|u_t(x_t, \mathbf{y} : \tau) - u_t(x_t, \mathbf{z} : \tau)| \leq |u_t(x_t, \mathbf{y}) - u_t(x_t, \mathbf{z})|,$$

where $\mathbf{y} : \tau$ denotes the sequential concatenation of $\mathbf{y}$ and τ and similarly for $\mathbf{z}$. In words, user realizes small utility gap if two outputs have more common tokens (words) in suffix.

For TLB problem, we propose an excessive optimism under the face of uncertainty (EOFUL) algorithm with regret $O(L\sqrt{T \log T})$,[3] by combining greedy decoding with standard LinUCB algorithm [178]. The analysis follows from several interesting ideas: (i) an introduction of level-k regret that internally occurs from k-th chosen token during the algorithm, (ii) fictitious extension of our algorithm's sequence to recursively propagate the regret into upper level, followed by (iii) regret de-

[3]The analysis requires an additional technical assumption, detailed in Section 8.3.2. Linear realizability and this assumption is completely removed in our results for TMAB, at the cost of the same context across the rounds.

composition and sum of squares regret analogous to the standard linear bandit analysis. For TMAB problem, we no longer have a global estimate θ that will help us estimating the performance of each intermediate token sequence. Instead, we present GreedyETC that step-by-step learns a desirable path to efficient token sequences by combining exploration-then-commit style of algorithm with greedy decoding, inducing the regret $O(nLT^{2/3}(\log T)^{1/3})$. Analysis requires a careful union bound over the concentration of desirable token sequences, since naively taking union bounds over every token sequence yields regret of $\Omega(n^L)$.[4] We remark that DDMC assumption plays a crucial role, enabling certain tractable structure in the utility function, albeit its naturallity and simplicity.[5] DDMC assumption is relaxed in Appendix G.2, and we also provide efficient algorithms for both settings with the weakened assumption by replacing the greedy decoding with lookahead decoding, another well-known decoding algorithm [179, 180].

Finally, we discuss several implications and applicability of our foundations in LLM decoding and alignment. For LLM alignment problem, we have a LLM equipped with a (known) utility function $v : \mathcal{V}^* \times \mathcal{V}^* \to \mathbb{R}_{\geq 0}$ and a misaligned utility represented by linear structure $\langle \theta, e(x_t, \mathbf{y}) \rangle$ with latent θ, where the eventual utility is a unknown convex combination over them. The objective is to efficiently learn θ and the convex combination parameter in a decoding time through user feedback while leaving the original LLM frozen to ensure that no further training occurs for the LLM. We show that this problem reduces to the TLB problem, enabling efficient learning via EOFUL under DDMC. We next consider a static version of our problem, where a query access to the sequence function u is available and we need to maximize u with polynomial queries. Inspired by our analysis for EOFUL and GreedyUCB, we prove that the greedy algorithm achieves the exact optimum with nLT queries.

[4] TMAB problem, in fact, has a close connection to the well known *bandit tree search* (BTS), *i.e.,* a decision version of BTS can be reduced to TMAB and vice versa. We discuss the detail in Appendix G.3.

[5] We also compare DDMC with smoothness assumption in BTS problem in Appendix G.3.

This immediately implies that greedy decoding is optimal under DDMC for LLM decoding problem when we view u as the probability of LLM generating the sequence. This gives an evidence on why greedy decoding is unreasonably efficient for various domains despite its simplicity [159, 181–183], (see [177] for comprehensive evaluation of greedy decoding).

8.2 Problem Setup

We formally define the general problem setup, which will subsume tokenized linear bandit (TLB) and tokenized multi-armed bandit (TMAB) problems. Let $\mathcal{V}$ be the set of all possible tokens with $|\mathcal{V}| = n$. For any set A, let A^* the collection of every possible finite sequences over A. Given a token sequence $\mathbf{y} \in \mathcal{V}^*$, we write $\mathbf{y}^{(i)} \in \mathcal{V}$ to denote its i-th token and $\mathbf{y}^{(i:j)} \in \mathcal{V}^*$ to denote the subsequence that contains every token from i-th token to j-th token including themselves for $j \geq i$ in $\mathbf{y}$.[6] We often write $\mathbf{y}^{(-i)}$ to denote the i-th token from the end of $\mathbf{y}$. There exists a special end-of-sequence $\texttt{EOS} \in \mathcal{V}$ token representing the end of a sequence. We call a token sequence $\mathbf{y}$ to be *complete* if its last token is $\texttt{EOS}$. For $\mathbf{y}, \mathbf{z} \in \mathcal{V}^*$, we write $\mathbf{y} : \mathbf{z}$ to denote the token sequence that concatenates $\mathbf{y}$ and $\mathbf{z}$ sequentially. We often write $\mathbf{y} = \emptyset : \mathbf{y}$.

We call the process of sequential token selection in a single round the *decoding* algorithm, and denote the process of observing the feedback at the end of the decoding and learning the sequence function by *learning* phase. Essentially, one can view the decoding process as an *online problem* to select each token, and learning process as an *online learning problem* to estimate the objective function. We write $|\mathbf{y}|$ to denote the length (interchangeably, depth or level) of the sequence $\mathbf{y}$. We consider a class of decoding algorithms with fixed maximal *depth* L that only outputs token sequence with length at most L. Equivalently, we assume $\mathcal{V}^*$ consists of token sequences with length up to L.

[6] We use boldface for sequences and non-boldface for a token.

We assume that the depth L is known in advance to the DM.[7] The depth plays a crucial role in our analysis and any algorithm's performance would depend on L that intrinsically captures the size of feasible token sequences. We refer to an algorithm's *level* l decision to denote its decision for the l-th token. Equivalently, if an algorithm reaches level $L-1$, then it observes nonzero utility only by appending `EOS`.[8]

Sequence utility and regret. At each round $t \in [T]$, a user arrives and writes a query $x_t \in \mathcal{V}^*$. Then, a decision maker (DM) needs to select each token sequentially to construct a sequence $\mathbf{y}_t \in \mathcal{V}^*$ to answer the user's query, and the DM observes a reward $r_t(\mathbf{y}_t)$ when it finalizes the sequence selection process and displays $\mathbf{y}$. In particular, the DM irrevocably commits to each token to construct the token sequence and cannot revert any previous decision to change the sequence entirely. The reward is defined by two factors: (i) a sequence (utility) function $u_t : \mathcal{V}^* \to \mathbb{R}_{\geq 0}$ that maps a token sequence to a nonnegative real value and (ii) a random noise over it. Note here that u_t can vary across the rounds, which captures the (possibly) different context provided to the DM given by the queries x_t over the rounds.[9] Formally, the random reward $r_t(\mathbf{y}_t)$ for the sequence $\mathbf{y}$ can be represented as

$$r_t(\mathbf{y}_t) = u_t(\mathbf{y}_t) + \eta_t,$$

where η_t is a random noise with $|\eta_t| \leq \sigma$.[10] The DM does not know the sequence function u in advance, but can only indirectly access it through the reward $r_t(\mathbf{y}_t)$. The goal of the DM is to maximize the user utility over the rounds T, *i.e.,* $\sum_{t\in[T]} r_t(\mathbf{y}_t)$. Equivalently, the objective is to compete with

[7] Practically, it can be prompted in the LLM or implicitly set in the LLM to ensure it does not generate an infinitely long output.

[8] In terms of LLM, this can be viewed as full probability being assigned to `EOS` at level L.

[9] An acute reader may easily observe that one cannot learn anything if u_t does not have any structural assumption. Thus, Section 8.3 allows efficient learning by having a linear structure, and Section 8.4 allows efficient learning my having $u_t = u$.

[10] We assume bounded noise for simplicity, but it easily extends to bounded sub-Gaussian noise as in the bandit literature.

the optimal complete sequence in hindsight: $\textsc{Opt} = \sum_{t=1}^{T} \textsc{Opt}_t$ where $\textsc{Opt}_t = \mathbb{E}\left[\max_{\mathbf{z}\in\mathcal{V}^*} r_t(\mathbf{z})\right]$. Then, the DM tries to minimize the expected cumulative (pseudo-) regret $\textsc{Reg} = \sum_{t\in[T]} \textsc{Reg}_t = \textsc{Opt} - \sum_{t\in[T]} \mathbb{E}\left[r_t(\mathbf{y}_t)\right]$.

`EOS` and monotonicity. Informally speaking, once the sequence meets `EOS`, it is natural to stop the decoding. To capture such behavior, we impose a natural assumption for the utility function based on the `EOS` token:

Assumption 8.2.1 (Monotonicity)**.** *The utility function u_t is monotone if (i) appending more* `EOS` *to a sequence that already has* `EOS` *at the end does not change the utility and (ii) appending non-*`EOS` *token to a sequence that already has* `EOS` *at the end only decreases the utility.*

In words, any reasonable algorithm would always not append any non-`EOS` token once after it has `EOS` as the last token, as it will always decrease the utility. Importantly, even though our sequence function could have nonzero value even without the `EOS` token at the end, we will *only* consider a class of algorithms that only ends with `EOS` token.

LLM decoding problem. Let us see the connection between our problem and *LLM decoding* problem and *LLM alignment* problem. Given an access to the logit probability of a LLM and a token sequence $\mathbf{y} \in \mathcal{V}^*$, suppose the sequence function u_t denotes the LLM's probability p of generating the output $\mathbf{y}$, *i.e.,* $u_t(\mathbf{y}) = p(\mathbf{y}) = \prod_{k=1}^{n} p(\mathbf{y}^{(k)}|\mathbf{y}^{(1:k-1)})$. Then, our objective can be framed as an online learning problem to learn the LLM's logit probability in an efficient manner.

A special case of this problem is a scenario where the DM has access to the value oracle of u, *e.g.,* knows the probability distribution, in advance. In this case, the objective boils down to the sequence function maximization problem [184–187] where duplicated insertion of the same element is allowed, or from the LLM's perspective, the *LLM decoding* problem of designing efficient decod-

ing algorithms [175, 188–190]. Correspondingly, our main results have several implication in these problems, which will be elaborated in Section 8.5.2.

LLM alignment problem. Further, one can consider a variant of our setup, termed *LLM alignment* problem, where we have a (frozen) LLM that gives us a probability distribution $p(\mathbf{y})$, but this LLM is (partially) misaligned with our objective function u. In this case, one standard approach would be to use the probability distribution $p(\mathbf{y})$ as a baseline for the decoding process, while learning the misaligned objective function u. If the misalignment allows specific structure, *e.g.*, $u_t(x_t, \mathbf{y}) = (1 - \gamma)p(x_t, \mathbf{y}) + \gamma f(x_t, \mathbf{y})$[11] for some latent function f capturing the misalignment, we indeed show that our results can be extended to efficiently handle such misalignment in Section 8.5.1.

8.3 Tokenized Linear Bandit

The general setting we formulated, however, does not allow any efficient learning if the utility function u_t indeed does not have structural correlation across the rounds. One natural framework to tackle such fundamental obstacle is to impose the *linear realizability* assumption in the online contextual learning literature. To this end, we introduce the following notion of embedding function:

Definition 8.3.1 (Embedding function)**.** A decoding algorithm has access to the embedding function $e : \mathcal{V}^* \times \mathcal{V}^* \to \mathbb{R}^d$ maps a token sequence to d-dimensional real vector.[12]

Assumption 8.3.2 (Linear realizability)**.** *There exists $\theta \in \mathbb{R}^d$ and embedding function $e(\cdot)$ such that the utility function can be represented by $u_t(x_t, \mathbf{y}) = \langle \theta, e(x_t, \mathbf{y}) \rangle$ for every query $x_t \in \mathcal{V}^*$ and token sequence $\mathbf{y} \in \mathcal{V}^*$.*

[11] In fact, setting $\gamma = 0$ reduces to the pure LLM decoding problem, whereas setting $\gamma = 1$ reduces to our general problem of learning the latent sequence function, or equivalently, learning a LLM's behavior.

[12] One can deem the embedding function as the semantic embedding that the LLM computes for $\mathbf{y}$ given a query x_t.

The DM has a query access to the embedding function during the decoding process.[13] Then, the DM's objective is to efficiently learn the latent parameter θ to maximize the utility. We assume that $\|\theta\| \leq 1$ and $\|e(\mathbf{z})\| \leq 1$.[14] We finally note that such linear approximation of the reward model often appear in the LLM alignment literature [191, 192].[15]

8.3.1 Diminishing distance with more commons (DDMC)

On the other hand, even if the utility function is well-structured with the linear realizability assumption, the intrinsic intricacy of TLB's tokenized nature imparts a fundamental hardness in maximizing the reward. Intuitively, if a (monotone) sequence function could have arbitrary structures, one small mistakes in the first token selection may incur irreparable results in the following tokens, *e.g.,* imagine there exists only a single token τ which incurs nonnegative for $u(\tau : \mathbf{y})$ for any $\mathbf{y}$. If the first token itself does not give much information about the following levels, one cannot fundamentally identify an efficient path in hindsight.[16] This is captured by the following proposition:

Proposition 8.3.3. *Under linearizability assumptinon with monotone utility function, any algorithm suffers the worst-case regret of* $\Omega(T(1 - 1/2^{L-1}))$.

Thus, we here introduce a fairly natural structural assumption on the sequence function $u(\cdot)$ to avoid such pessimistic scenario, dubbed *diminishing distance with more commons*.

Assumption 8.3.4 (DDMC)**.** *A sequence function* $u(\cdot)$ *has* diminishing distance with more commons

[13] Our algorithm calls the embedding function $O(nLT)$ time.

[14] This is for ease of exposition to exclude cumbersome details, but all results would easily generalize under any bounded scenario.

[15] One might find this partly related to linear representation hypothesis [193], stating that concepts are encoded as linear directions in the model's embedding space.

[16] In fact, due to the online learning nature of our problem, there always exists a small probability that any algorithm makes an error in the first level.

(DDMC) if

$$u(x_t, \mathbf{y} : \tau) - u(x_t, \mathbf{z} : \tau) \leq u(x_t, \mathbf{y}) - u(x_t, \mathbf{z}),$$

for any two finite sequences $\mathbf{y}, \mathbf{z}$ *with the same length* $|\mathbf{y}| = |\mathbf{z}|$*, any token* $\tau \in \mathcal{V}$*, and any query* x_t *such that LHS is nonnegative, and further if* $\mathbf{y}$ *and* $\mathbf{z}$ *has a single* `EOS` *at their ends, then analogous inequality holds by appending* τ *right before the* `EOS`. [17]

In other words, adding common tokens at the end of two sequences with the same length only decrease the utility gap between them.

DDMC shares very similar intuition from the widely adopted *submodularity* assumptions on set function [194], or recently sequence function [187], which characterizes diminishing returns. However, they crucially differ and cannot be implied by each other, due to two key differences: (1) submodularity is about diminishing *return* (i.e., the difference between two function values) whereas DDMC is about diminishing *distance* (i.e., absolute value of the difference); (2) more subtly, DDMC is about the difference of two different sequences of the same length with an added common token, whereas submodularity is about the difference of two different sets that one includes another with an added common element. In some sense, DDMC is the complement of submodularity.

Discussion 8.3.5. From the practical perspective for LLM, this is innocuous since if two generated outputs shares common tokens suffix-wise, we can expect that they give the user similar experiences as long as we add more commons.

[17] In Appendix G.2, we provide a relaxed version of DDMC assumption and provide corresponding efficient algorithms.

8.3.2 Excessive Optimism under the Face of Uncertainty

We present our algorithm, named *excessive optimism under the face of uncertainty* for tokenized linear bandits (EOFUL), whose pseudocode is presented in Algorithm 2. It has *two optimistic* features. First, it optimistically decodes each token in a greedy manner, expecting that this would lead to a good utility in following decisions. Second, it optimistically estimates each token's utility by constructing a confidence ball C_t, and computes the best possible utility each token can gain from parameters in C_t.

Our construction of the confidence ball is analogous to the standard LinUCB [178, 195] algorithm for linear bandits.

Let $\mathbf{y}_t$ be the chosen sequence at round t by Algorithm 2. We recursively define the following matrices for $t \in [T]$:

$$\Sigma_t = \Sigma_{t-1} + \mathbf{y}_t \mathbf{y}_t^\top,$$

and let $\Sigma_1 = \lambda I$ for $d \times d$ identity matrix I. Further, we define

$$\beta_t = \sigma^2 \left(2 + 4d \log \left(1 + \frac{tL}{d} \right) + 8 \log \frac{4}{\delta} \right).$$

For $t \geq 2$, the center of our confidence ball will be:[18]

$$\hat{\theta}_t = \Sigma_t^{-1} \sum_{i=1}^{t-1} r_t(x_t, \mathbf{y}_t) \mathbf{y}_t, \tag{8.1}$$

and $\hat{\theta}_1 = \vec{0}$, the d-dimensional zero vector.[19]

[18] Remark that this is a solution of the ridge regression given the reward feedback.

[19] For efficient computation, one may use the rarely-switching version of OFUL algorithm in [178], instead of the stan-

Now, our confidence ball C_t to be used at round t is defined as follows:

$$C_t = \{\vartheta : (\vartheta - \hat{\theta}_t)^\top \Sigma_t (\vartheta - \hat{\theta}_t) \leq \beta_t\}, \tag{8.2}$$

and $C_1 = \{\vec{0}\}$. We choose $\lambda = \Theta(1)$.

ALGORITHM 2: EOFUL

Input: Token set $\mathcal{V}$, linear contextual bandit algorithm ALG
Initialize $\mathbf{y} \leftarrow \emptyset$, $\hat{\theta}_1 \leftarrow \vec{0}$, $C_1 = \{\hat{\theta}\}$, $\mathbf{y} = \emptyset$
for $t = 1, 2, \ldots, T$ **do**
  User arrives and submits query x_t
  for $k = 1, 2, \ldots, L-1$ **do**
    Compute $\tau^* = \mathrm{argmax}_{\tau \in \mathcal{V}, \theta \in C_t} \langle \theta, e(x_t, \mathbf{y} : \tau) \rangle$
    Set $\mathbf{y} \leftarrow \mathbf{y} : \tau$
    if $\tau =$ `EOS` **then** Break;
  Submit $\mathbf{y}$ and observe reward r_t
  Compute $\hat{\theta}_{t+1}$ by (8.1) and C_{t+1} by (8.2)

For the analysis, we impose the following technical assumption:

Assumption 8.3.6. *For any complete sequence* $\mathbf{y}$ *with* $|\mathbf{y}| = k$ *and any subsequence* $\mathbf{y}^{(1:l)}$ *with* $l \leq k$*, there exists a* $c \geq 0$ *such that* $\left\|\mathbf{y}^{(1:l)}\right\|_{\Sigma_t} \leq c \left\|\mathbf{y}\right\|_{\Sigma_t}$ *in Mahalanobis norm.*[20]

That is, each entry of the subsequence of a complete sequence is not too larger than that of the complete sequence with respect to Σ_t. Both structural assumptions 8.3.2 and 8.3.6 are removed in Section 8.4 for TMAB.

Discussion 8.3.7. This assumption is innocuous in a sense that c can be sufficiently large up to $o(\sqrt{T})$ to obtain sublinear regret. On the other hand, this is less innocuous since it depends on the algorithm's construction of Σ_t. In practice, one of the typical usage for LLM is to either (i) have similar queries based on their daily lives, or (ii) ask queries in already existing sessions (in-context) to expand the

dard OFUL.

[20] Mahalanobis norm is defined by $\|x\|_A = \sqrt{x^\top A x}$.

discussion. In these scenarios, we expect the queries and the responses thereby to be not too much different from each other.

Discussion 8.3.8. On the other hand, if $\mathbf{y}_i^{(1:l)}$ and $\mathbf{y}_t$ does not have any structural dependency, any algorithm would not be able to fundamentally learn certain coordinate of θ. For instance, suppose any complete sequence's embedding only has nonzero values in the last $d/2$ dimension, and the first $d/2$ dimension for any incomplete sequence. Then, any algorithm could not learn the first $d/2$ dimension, implying that any internal node does not provide information on an efficient path to the complete sequence. Relaxing this assumption in a nuanced way remains a major open problem.

Theorem 8.3.9. *With probability at least* $1-\delta$*, EOFUL has regret* $O(cL\sqrt{dT(\log T + \log(\frac{1}{\delta}))}$*, and thus plugging* $\delta = 1/T$ *yields* $\textsc{Reg} = O(cL\sqrt{dT\log T})$.

Overall, our proof proceeds as follows:

Step 1 (Length equalization). We first equalize $\mathbf{o}_t$ and $\mathbf{y}_t$ to have the same lengths in a way that it does not hurt any following analysis.

Step 2 (Level-k regret). We define level k regret as the regret occurred at depth k of the decoding process that will play crucial role in the analysis to recursively transform the regret into upper level regrets.

Step 3 (Fictitious extension). Then we introduce a fictitious extension of $\mathbf{y}_t^{(1:k-1)}$ to $\mathbf{f}_t^{(1:k)} = \mathbf{y}_t^{(1:k-1)} : \mathbf{o}_t^{(k)}$, and show how the level-k regret of the fictitious extension is upper bounded by level-(k-1) regret at $\mathbf{y}_t^{(1:k-1)}$.

Step 4 (Regret decomposition). We relate the level-(k-1) regret of the algorithm's choice $\mathbf{y}_t$ and level-k regret of the fictitious extension using DDMC assumption, which upper bounds our algorithm's level-k regret due to the optimistic choice of the algorithm. Then, we telescope over $k, \dots, 1$ to express

the entire regret as a function of level-1 regret plus some estimation errors occurred for each prefix of $\mathbf{y}_t$ during the decoding.

Step 5 (Sum of squares regret) Finally, we compute the function of level-1 regret and the estimation error by following the standard sum of squares regret bound analysis of linear contextual bandit algorithm, where readers familiar with the standard linear bandit analysis may easily follow.

8.4 Tokenized Multi-armed Bandit

The linear realizability assumption is essential for the contextual setting to derive any efficient algorithm. A natural follow-up question is whether one could obtain an efficient algorithm without linear realizability, if the context does not change and the utility function $u_t = u$ remains the same over the rounds.[21] In this case, the DM's problem can be seen as a variant of the standard multi-armed bandit problem where it needs to make an (online) decision for the next token irrevocably to construct a complete sequence. Thus, we dub this problem *tokenized multi-armed bandit* (TMAB).

We first argue that even in this simple setting, our problem admits a pessimistic lower bound when the sequence function does not have any structural assumption beyond monotonicity. Then, we again introduce DDMC and present an algorithm with regret $O(nLT^{2/3}(\log T)^{1/3})$. We present an intimate connection between TMAB and the bandit-based tree search (BTS) problem [176] in Appendix G.3.

Exponential lower bound. If the query is fixed, our objective boils down to efficiently learn the latent utility function u. Without any structural assumption on it, however, one can easily see that it is equivalent to learn the reward distribution of every subsequence $\mathbf{y} \in \mathcal{V}_L^*$. This is because each

[21] We omit the context x_t in the inputs of u since the context remains the same.

subsequence $\mathbf{y}$ can be deemed as a single arm in the standard multi-armed bandit problem which admits regret lower bound of $\Omega(\sqrt{KT})$ where K is the number of arms. This is formally stated as follows:

Proposition 8.4.1. *For TMAB problem with monotone utility function, any algorithm has worst-case regret lower bound of* $\Omega(\min(n^L\sqrt{T}, T))$.

DDMC and GreedyETC. The pessimistic exponential dependency on depth L naturally motivates the use of DDMC Assumption 8.3.4. Unlike the TLB problem, however, there exists no global latent parameter θ that we could use throughout the decoding process, but they can genuinely arbitrary beyond DDMC (and monotonicity). Thus we use a rather direct approach to explore each level sufficiently, and commit to the seemingly optimal token at each level then move on to the next level.

ALGORITHM 3: GreedyETC

Input: Token set $\mathcal{V}$, exploration parameter N
Initialize $\mathbf{y} \leftarrow \emptyset$
for $k = 1, 2, \ldots, L$ **do**
  if $|\mathbf{y}| = L - 1$ **then**
    $\mathbf{y} \leftarrow \mathbf{y} : \texttt{EOS}$
    Break
  for $t \in \mathcal{V}$ **do**
    Submit $\mathbf{y} : t : \texttt{EOS}$ for N times and observe rewards
    Set $\bar{r}(\mathbf{y} : t)$ be the averaged cumulative rewards
  Select $\tau^* = \operatorname{argmax}_{t \in \mathcal{V}} \bar{r}(\mathbf{y} : \tau)$
  Set $\mathbf{y} \leftarrow \mathbf{y} : \tau^*$
  if $\tau^* = \texttt{EOS}$ **then**
    Break
For the rest of the rounds, submit $\mathbf{y}$

We prove that under DDMC assumption, a natural combination of greedy decoding with exploration-then-commit (ETC) style of algorithm, as presented in Algorithm 3, yields sublinear regret w.r.t. T with linear dependency on d:

Theorem 8.4.2. *Under Assumption 8.2.1 and 8.3.4, Algorithm 3 achieves regret* $O(nLT^{2/3}(\log T)^{1/3})$.

One subtle part is that, if one naively construct the *clean* event that the rewards for each sequence are concentrated to its expectation and take union bound over all of them, this essentially gives the regret of $\Omega(n^L)$, which is undesirable. Instead, we only take the union bound along the path that our algorithm is selected, which enables us to recover the linear dependency on L. Then, we obtain a recursive inequality relating the regret at level k with that at level $k-1$ by considering fictitious extension as defined in the proof of Theorem 8.3.9, and conclude that the regret remains small with high probability.

8.5 Applications and Experiments

8.5.1 Application: LLM alignment

One interesting application of our results is the LLM alignment problem. In this problem, we want to generate aligned outputs given access to a frozen LLM that cannot be further fine-tuned. Consider a LLM parameterized by a token distribution $p : \mathcal{V}^* \times \mathcal{V}^* \to [0,1]$ that maps any token sequence given a user's query to a probability that the LLM generates this sequence. The user's utility given token sequence $\mathbf{y}$ is misaligned from the LLM such that $u(x_t, \mathbf{y}) = \gamma v(p(x_t, \mathbf{y})) + (1-\gamma) f(x_t, \mathbf{y})$, where v is a monotone function that maps probability to reward and f is the latent function to be learned that marginalizes the misaligned part of the LLM.

Such convex combination between multiple rewards typically appear in the LLM alignment literature [173, 196–198], when one wants to align with multiple human preferences by well combining multiple models or parameters. Our utility function as the convex combination over v and f can also be interpreted in a similar spirit. One typical example for the function v would be the log probability or perplexity obtained from the frozen LLM, or more explicit user utility such as classification accuracy

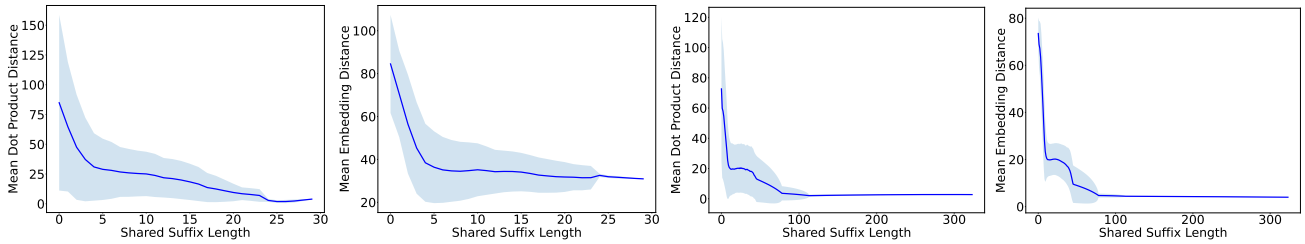


Figure 8.1: The first two show distance gap under TruthfulQA and the next two is under HH-RLHF, with distance function d_1 and d_2.

or factuality [199]. We assume the DM knows the probability to utility mapping function v.

If we assume f is linearly realizable, *i.e.,* there exists an embedding function $e : \mathcal{V}^* \times \mathcal{V}^* \to \mathbb{R}^d$ such that $f(\cdot) = \langle \theta, e(x_t, \mathbf{y}_t) \rangle$, we can reduce the problem of learning misaligned function f to the TLB problem. Formally, we construct $\theta' \in \mathbb{R}^{d+1}$ as a vector that concatenates θ and 1 in $d+1$-th dimension, *i.e.,* $\theta' = [\theta, 1]$. Then we can rewrite the utility as

$$\begin{aligned} u(x_t, \mathbf{y}) &= \langle \theta, (1-\gamma) e(x_t, \mathbf{y}_t) \rangle + \gamma v(p(x_t, \mathbf{y})) \\ &= \langle \theta', [(1-\gamma) e(x_t, \mathbf{y}), \gamma v(p(x_t, \mathbf{y}))] \rangle \,. \end{aligned}$$

This is exactly the feedback structure in our TLB problem, except that the DM knows the last coordinate of the latent parameter θ, which only makes the problem easier. Correspondingly, EOFUL enjoys the same regret bound in Theorem 8.3.9 for this problem.[22]

8.5.2 Implication: LLM decoding

Our results and analysis therein have a few interesting implications in (i) the sequence function maximization problem [184–186, 200] and (ii) the problem of designing efficient decoding algorithms

[22]Note that our formulation differs from standard KL-divergence based problem [192]. One might integrate the KL divergence in our problem too by redefining the function over the distribution rather than a fixed sequence, however, it's not clear to formulate in the language of multi-armed bandit problem. This remains an interesting future direction.

for text generation in language model [175, 188–190], if we consider a special case when the function u is known to the DM and consider a static version with a single round.

First, this is exactly a sequence function maximization problem to maximize u given query access to a value oracle that outputs $u(x_t, \mathbf{y})$ given a sequence $\mathbf{y}$. One common objective here is to obtain an approximate algorithm with polynomial query complexity. If u satisfies a notion of monotonicity and sequence-submodularity [200], it is known that the greedy algorithm has $1 - 1/e$ approximation.[23] On the other hand, our result implies that greedy algorithm is exactly optimal under monotone DDMC sequence function, formally stated as follows:[24]

Theorem 8.5.1. *For any monotone sequence function u, the greedy algorithm is optimal under DDMC and monotonicity.*

From the LLM perspective, this result has one interesting corollary. Given a fixed query x, consider a LLM parameterized by $p : \mathcal{V}^* \times \mathcal{V}^* \to [0, 1]$. Then, Theorem 8.5.1 implies that if p is monotone DDMC function, the greedy decoding can exactly output the token sequence that maximizes the logit probability. This supports the unreasonable effectiveness of the greedy decoding, despite its simplicity, justifying why it works well and widely adopted in the literature [177, 181, 182], *e.g.,* in Google's Gemini report [159], even though it is only known as an approximate maximum-a-posterior decoding algorithm [183].

On the other hand, greedy decoding is often dominated by sampling methods such as nucleus sampling [201] or self-consistency [202] for various tasks, *e.g.,* [177] find that vanilla multinomial sampling can be better than greedy for open-ended creative generation. We conjecture that such phe-

[23] It also admits the APX-hardness with $1-1/e$ by the hardness of submodular set-function maximization from Maximum Coverage, unless P=NP.

[24] We remark that our setup is slightly different from standard sequence maximization problems as we allow the selection of the same element over different position, and the special `EOS` token governs the end of the sequence by naturally encoding it in the monotonicity of the utility function.

nomenon largely depends on how the sequence function generated by LLM, *i.e.,* the logit probabilities, would look like and thereby the type of the tasks. Investigating other regimes that another decoding algorithm provably outperforms greedy remains an intriguing open question.

8.5.3 Validating DDMC

We finally provide some experimental results on whether DDMC indeed holds for certain tasks.

Setup. For the embedding function e, we obtain the embedding from Llama3-8B-Instruct model [203]. For the token sequences, to validate our DDMC in real-world datasets, we use TruthfulQA dataset [199] and HH-RLHF dataset [164],[25] by concatenating each query and answer pair.

For each item in the dataset, we group every (query:answer) pair by number of common tokens in its suffix. For instance, image the following two token sequences:

$$\mathbf{x} = (a, b, c, d, e, f),\ \mathbf{y} = (a, x, y, d, e, f).$$

Then, we can group each subsequence as the following:

- 0 common suffix: $(\mathbf{x}^{(1:2)}, \mathbf{y}^{(1:2)}), (\mathbf{x}^{(1:3)}, \mathbf{y}^{(1:3)})$
- 1 common suffix: $(\mathbf{x}^{(1:1)}, \mathbf{y}^{(1:1)}), (\mathbf{x}^{(1:4)}, \mathbf{y}^{(1:4)})$
- 2 common suffix: $(\mathbf{x}^{(1:5)}, \mathbf{y}^{(1:5)})$
- 3 common suffix: $(\mathbf{x}^{(1:6)}, \mathbf{y}^{(1:6)})$

We evaluate each distance $d(\mathbf{x}^{(1:k)}, \mathbf{y}^{(1:k)})$ in each group and obtain the average and variance.

[25] Both datasets are standard for LLM alignment, where TruthfulQA validates truthfulness and hallucination and HH-RLHF evaluates helpfulness and harmlessness. In HH-RLHF, data are organized in (prompt:response) format. For simplicity, we reuse "query" to refer to "prompt" and "answer" to refer to "response".

For the distance functions, we evaluate two candidates:

$$d_1(x, y) = \langle \theta, \mathbf{x} - \mathbf{y} \rangle$$

$$d_2(x, y) = \|\mathbf{x} - \mathbf{y}\|_2 ,$$

where $\theta = \vec{1}_{4096}$, *i.e.,* 4096-dimensional one vector. The first distance function d_1 is consistent with our DDMC assumption, and it in fact can be viewed as a ℓ_1 distance between two vectors. Thus, the second distance function d_2 can be viewed as a generalization of d_1 into ℓ_2 distance.

Result. The first two figures in Figure 8.1 refers to the TruthfulQA, and the next two denotes the HH-RLHF. Each experiment is presented for distance function d_1 and then d_2. Overall, we observe a tendency to have decreasing distance gap as the number of common tokens in suffices increases. This holds for both the distance function d_1 and d_2. Interestingly, the decreasing curvature tends to be stronger when the number of common tokens are rather small. Further, the decreasing gap seems to be more drastic under HH-RLHF, suggesting that different tasks might possess different structures in their sequence function, *i.e.,* logit probability structure. Overall, despite verified in an average sense, this justifies the DDMC assumption, which we believe could be largely useful for following works.

Appendix A: Delegated Choice with Combinatorial Constraints

A.1 Further Preliminaries on Matroids

We further present several results on matroids, which will be useful throughout the paper in proving main results. First the following lemma states that the restriction and contraction over a matroid induces a matroid.

Lemma A.1.1 ([204]). *Given a matroid $\mathcal{M} = (U, \mathcal{I})$, for any $V \subseteq U$, the restriction $\mathcal{M}|V$ and contraction $\mathcal{M}/V$ are matroids.*

ALGORITHM 4: Greedy

Input: Set system $\mathcal{M} = ([n], \mathcal{I})$ where weight of i is $w_i \geq 0$ for $i \in [n]$
Sort the elements of the set system so that $w_1 \geq w_2 \geq \ldots \geq w_n$.
for $i = 1, 2, \ldots, n$ **do**
  $S \leftarrow \emptyset$
  if $S \cup \{i\} \in \mathcal{I}$ **then**
    $S \leftarrow S \cup \{i\}$.
return S

The following is a well-known result on the matroid, stating that the greedy algorithm is optimal under matroid constraint, and even the matroid is necessary for greedy to be optimal.

Lemma A.1.2 ([204]). *Let $\mathcal{I}$ be a collection of subsets of U. Then, $(U, \mathcal{I})$ is a matroid if and only if the following properties hold:*

1. *$\emptyset \in \mathcal{I}$.*
2. *If $I \in \mathcal{I}$ and $I' \subset I$ then $I' \in \mathcal{I}$.*

3. *For all weight functions $w : U \mapsto \mathbb{R}$, the greedy (Algorithm 4) produces a maximal member of $\mathcal{I}$ with maximum weights.*

Finally, a union over matroid, where the corresponding independent set is defined as union over each corresponding matroid, is remained to be a matroid.

Lemma A.1.3 ([204]). *For $i \in [k]$, let $\mathcal{M}_i = (U_i, \mathcal{I}_i)$ be a matroid. Then, the matroid union defined by*

$$\mathcal{M} = \cup_{i\in[k]}\mathcal{M}_i = (\cup_{i\in[k]}U_i, \mathcal{I} = \{\cup_{i\in[k]}I_i : I_i \in \mathcal{I}_i\}),$$

is also a matroid.

A.2 Implementation of Agent-aware Mechanisms

First, the following is an analogue of Theorem 1.6.1 for agent-aware mechanisms.

Proposition A.2.1 (Restatement of Proposition 1.2.7; Formal). *Any α-competitive order-hinted non-adaptive algorithm for matroid prophet inequality implements an α-approximate agent-aware mechanism for the matroid delegated choice.*

It is worth noting that the analogues of matroid-constrained/boundary-randomized non-adaptive version also carries over. The arguments are essentially cumbersome variants of the proofs for agent-oblivious mechanism, so we omit the details.

This has two implications: (i) implementability of the order-aware algorithm for matroid prophet inequality in the *deterministic agent setting* and (ii) implementability of non-adaptive algorithm for matroid prophet secretary in the *ex-ante indifferent agent* setting.

Corollary A.2.2. *Consider a* deterministic agent *setting,* i.e., *the matroid delegated choice problem with deterministic agent's utilities where Y_i has a singleton support for every $i \in U$. In this setting, any α-competitive order-aware non-adaptive algorithm for matroid prophet inequality implements an α-approximate agent-aware mechanism for the matroid delegated choice.*

Corollary A.2.3. *Consider an* ex-ante indifferent agent *setting,* i.e., *matroid delegated choice problem with the same agent's utility distribution,* i.e., *$Y_i \sim Y_j$ for every $i, j \in U$. In this setting, any α-competitive non-adaptive algorithm for matroid prophet secretary implements an α-approximate agent-aware mechanism for the matroid delegated choice.*[1]

All the proofs can be found in Appendix A.4. An interesting open question in this context is, whether there exists a constant-approximate non-adaptive algorithm for matroid secretary, which has direct implication on matroid delegated choice with ex-ante indifferent agents.

A.3 Some Probability Tools and Inequalities

In this section, we present some probability tools that are used throughout the proof of the main results. We start with the following well-known max-min inequality.

Lemma A.3.1 (Max-min inequality)**.** *Given any set Z and W, for any function $f : Z \times W \mapsto \mathbb{R}$, it follows that*

$$\sup_{z \in Z} \inf_{w \in W} f(z, w) \leq \inf_{w \in W} \sup_{z \in Z} f(z, w).$$

If the function f satisfies certain properties, one can in fact observe an equality above. This is

[1] This corollary suggests the importance of constructing non-adaptive algorithms for (matroid) prophet secretary, which has not received any attention in the literature, to the best of the authors' knowledge.

characterized by the following minimax theorem due to [205], which in fact is a special case with a linear function of the original theorem.

Lemma A.3.2 (Minimax theorem)**.** *Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be compact convex sets. If $f : X \times Y \to \mathbb{R}$ is a continuous function that is concave-convex,* i.e., *$f(\cdot, y) : X \to \mathbb{R}$ is concave for fixed y and $f(x, \cdot) \to \mathbb{R}$ is convex for fixed x, then we have*

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

As a special case, if $A \in \mathbb{R}^{n \times m}$ is a finite matrix, and Then, we have

$$\max_{x \in X} \min_{y \in Y} x^\intercal A y = \min_{y \in Y} \max_{x \in X} x^\intercal A y.$$

The following lemma is a folklore technique in the prophet inequality and online stochastic optimization stating that it is without loss of generality (in terms of the approximation ratio) to consider problem instances such that the distribtions are continuous and density functions are uniformly bounded above by a constant.

Lemma A.3.3 (Continuous reduction for prophet inequality)**.** *Let* ALG$'$ *be α-competitive (non-adaptive) algorithm for the prophet inequality under the set system $\mathcal{M} = (U, \mathcal{I})$, under the problem instance where the random variables' distributions are continuous and density functions are uniformly bounded above by a constant. Then, for any $\varepsilon > 0$ there exists $(\alpha - \varepsilon)$-competitive (non-adaptive) algorithm* ALG *for arbitrary distribution.*

Proof. Similar to Theorem A.4.4, we add a noise to convert the distribution to one that is absolutely continuous with a bounded loss in the competitive ratio. Consider a Gaussian noise vector

$\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ where each ε_i is drawn from $\mathcal{N}(0, \sigma)$. The parameter σ will be defined shortly. Given the random variables $(W_i)_{i \in U}$, define $W'_i = W_i + |\, \varepsilon_i \,|$. By Lemma A.3.4, the probability distribution of W'_i is absolutely continuous for $i \in U$ and its density is uniformly bounded above.

Let I' be the random set denoting the selected set under ALG$'$, and I for ALG. Let Θ be the problem instance $(\mathcal{M}, W, T)$ and Θ' be $(\mathcal{M}, W', T)$. We write $x(I, \Theta)$ to denote the weights of I under Θ. Since ALG$'$ is α-competitive on Θ', we have

$$\frac{\mathbb{E}\,[\, x(I', \Theta')\,]}{\mathbb{E}\,[\, \max_{J \in \mathcal{I}} x(J, \Theta')\,]} \geq \alpha.$$

Now, we construct ALG as follows.

- It first independently samples ε_i from $\mathcal{N}(0, \sigma)$ for each $i \in U$.
- If element i arrives with W_i, it feeds W'_i to ALG$'$ and accepts the element iff ALG$'$ accepts i.

First, we have

$$\mathbb{E}\left[\max_{J \in \mathcal{I}} x(J, \Theta)\right] \leq \mathbb{E}\left[\max_{J \in \mathcal{I}} x(J, \Theta')\right]. \tag{A.1}$$

Meanwhile, we also have

$$\begin{aligned} x(I, \Theta) = x(I', \Theta') - \sum_{i \in I'} |\, \varepsilon_i \,| &= \sum_{i \in I_c} W'_i - \sum_{i \in I_c} |\, \varepsilon_i \,| \\ &\geq \sum_{i \in I_c} W'_i - \left(\sum_{i \in U} |\, \varepsilon_i \,|\right). \end{aligned}$$

Thus, we obtain

$$\mathbb{E}\,[\,x(I,\Theta)\,] \geq \mathbb{E}\,[\,x(I',\Theta')\,] - O(n\sigma). \tag{A.2}$$

Combining (A.1) and (A.2), we obtain the desired result.

Finally, it is easy to see that if ALG$'$ is non-adaptive then ALG is also non-adaptive. This is because our construction of ALG can be replaced by

- It first independently samples ε_i from $\mathcal{N}(0,\sigma)$ for each $i \in U$.
- Call ALG$'$ on W_i' to obtain thresholds τ'.
- Accept element i if it is feasible and $W_i \geq \tau_i' - \varepsilon_i = \tau_i$.

The last step guarantees that this is indeed a non-adaptive algorithm. a

The following elementary lemma is required to prove Lemma A.3.3.

Lemma A.3.4 (Density of noisy distribution)**.** *Given random variables X and Y, denote their probability distribution functions F_X and F_Y, and let f_x be the density function of X. Then, $X+Y$ has a density function,* i.e., *it is absolutely continuous. Further, its density can be expressed by*

$$f_{X+Y}(z) = \mathbb{E}\,[\,f_X(z-Y)\,].$$

Proof. It suffices to show that integrating RHS gives $F_{X+Y}(z)$. Note that the result then follows from

the fundamental theorem of calculus. Observe that

$$\begin{aligned}
\int_{-\infty}^{z} f_{X+Y}(w)dw &= \int_{-\infty}^{z} \int_{\mathbb{R}} f_X(w-y)\ \mathtt{d}F_Y(y)\ \mathtt{d}w \\
&= \int_{\mathbb{R}} \int_{-\infty}^{z} f_X(w-y)\ \mathtt{d}w\ \mathtt{d}F_Y(y) && \text{(Fubini's theorem)} \\
&= \int_{\mathbb{R}} \int_{-\infty}^{z-y} f_X(x)\ \mathtt{d}x\ \mathtt{d}F_Y(y) && \text{(Change of variables } x = w - y\text{)} \\
&= \int_{\mathbb{R}} F_X(z-y)\ \mathtt{d}F_Y(y).
\end{aligned}$$

Note further that this is exactly the expectation of $F_X(z-y)$, so we obtain

$$\begin{aligned}
\int_{-\infty}^{z} f_{X+Y}(w)dw = \mathbb{E}\left[\, F_X(z-Y) \,\right] &= \mathbb{E}_Y\left[\Pr\left[\, X \leq z - Y \,\right]\right] \\
&= \mathbb{E}_Y\left[\mathbb{E}_X\left[\mathbf{1}\left\{\, X + Y \leq z \,\right\}\right]\right] \\
&= \Pr\left[\, X + Y \leq z \,\right] \\
&= F_{X+Y}(z),
\end{aligned}$$

which completes the proof. a

Finally, the lemma below asserts that a random variable's expectation tailored to a specific realization of the random variable is maximized when it is conditioned to be its tail values.

Lemma A.3.5. *Let X be a random variable with a continuous distribution. Let E be an event with probability p. Set τ such that $\Pr\left[\, X \geq \tau \,\right] = p$. Then, we have*

$$\mathbb{E}\left[\, X \cdot \mathbf{1}\left\{\, X \geq \tau \,\right\}\right] \geq \mathbb{E}\left[\, X \cdot \mathbf{1}\left\{\, E \,\right\}\right].$$

Proof. Define E' as the event that $X \geq \tau$. We claim that

$$\Pr[X > x \text{ and } E] \leq \Pr[X > x \text{ and } E']$$

This is because if $x \geq \tau$, then $X > x$ and E' is the same event as $X > x$ and if $x < \tau$, then $X > x$ and E' is the same event as E' and $\Pr[E'] = \Pr[E]$. It follows that

$$\begin{aligned}
\mathbb{E}[X \cdot \mathbf{1}\{E\}] &= \int_{x>0} \Pr[X \cdot \mathbf{1}\{E\} > x]dx \\
&= \int_{x>0} \Pr[X > x \text{ and } E]dx \\
&\leq \int_{x>0} \Pr[X \geq x \text{ and } E']dx \\
&= \mathbb{E}[X \cdot \mathbf{1}\{E'\}]
\end{aligned}$$

a

A.4 Proofs

A.4.1 Proof of Theorem 1.5.1

To prove the theorem, we require the following technical lemma on concentration for self-bounding function.

Lemma A.4.1 ([51]). *Let $X = (X_i)_{i \in U}$ be a finite family of independent random variables. Let $Z = g(X)$ be a nonnegative bounded measurable function of X, and define $X^{(i)} = (X_j)_{j \in U \setminus \{i\}}$. Suppose that there exists some measurable function $Z^{(i)}$ of $X^{(i)}$ such that Z is self-bounding: (i) (1-Lipschitz) $0 \leq Z - Z^{(i)} \leq 1$ and (ii) (self-bounding) $\sum_{i \in U} Z - Z^{(i)} \leq Z$. Then, the following holds*

for Z.

$$\Pr[\, Z \leq \mathbb{E}[\, Z\,] - t\,] \leq \exp\left(-\frac{t^2}{2\mathbb{E}[\, Z\,]}\right).$$

Proof of Theorem 1.5.1. Before presenting our mechanism, we first trim the set system by removing element i if there exists no feasible set I that contains i. By doing so, we can get rid of any redundant element in U, thereby removing trivial arguments to deal with it. Note that this procedure obviously does not change any approximation factor.

Given a threshold $\tau_t \geq 0$, the *tail* mechanism $\textsc{Mec}_t$ is defined by the following rules: $\textsc{Mec}_t$ accepts the proposed set $I \in \mathcal{I}$ if there exists at least one element $i \in I$ satisfying $X_i \geq \tau_t$. Given another threshold τ_c, the *core* mechanism $\textsc{Mec}_c$ accepts the proposed set $I \in \mathcal{I}$ if the summation of the principal's utility $\sum_{i \in I} X_i \geq \tau_c$.

Importantly, we will show that there exists a pair (τ_t, τ_c) with a good approximation ratio. To this end, we first define some notations. Let τ_t be a nonnegative real value such that $\prod_{i \in U} \Pr[\, X_i < \tau_t\,] = \alpha$ for some $\alpha \in (0, 1)$, which will be determined later. Note that τ_t is a function of α. Define $p_i = \Pr[\, X_i \geq \tau\,]$ and $v_i = \mathbb{E}[\, X_i | X_i \geq \tau_t\,]$. Further, let Z denote the random set defined by $Z = \{i \in U : X_i < \tau_t\}$, and let E be the event that $Z = U$, *i.e.,* $X_i < \tau_t$ for every $i \in U$. Denote the optimal feasible set by I^*, *i.e.,* $X_{I^*} \geq X_I$ for every $I \in \mathcal{I}$. If there are multiple such optimal sets, arbitrarily pick one of them.

We first show that under $\textsc{Mec}_t$ with τ_t, the following inequality holds

$$\mathbb{E}[\, \textsc{Mec}_t \mid E^c\,]\Pr[\, E^c\,] \geq \alpha \sum_{i \in U} p_i \cdot v_i$$

To this end, let L_i be the event such that $X_i \geq \tau$ while $X_j < \tau$ for every $j \neq i$. Then, we observe that $\cup_{i \in U} L_i \subset E^c$, and $L_i \cap L_j$ has measure zero for every $i \neq j \in U$. Hence, we can decompose $\mathbb{E}\left[\,\textsc{Mec}_t \mid E^c\,\right]$ as follows.

$$\begin{aligned}
\mathbb{E}\left[\,\textsc{Mec}_t \mid E^c\,\right]\Pr\left[\,E^c\,\right] &\geq \sum_{i \in U} \mathbb{E}\left[\,\textsc{Mec}_t \mid L_i, E^c\,\right]\Pr\left[\,L_i \mid E^c\,\right] \\
&= \sum_{i \in U} \mathbb{E}\left[\,\textsc{Mec}_t \mid L_i\,\right]\Pr\left[\,L_i\,\right] \\
&= \sum_{i \in U} \mathbb{E}\left[\,\textsc{Mec}_t \mid L_i\,\right]\Pr\left[\,X_i \geq \tau_t\,\right]\Pr\left[\,X_j < \tau_t, j \neq i, \forall j \in U\,\right] \\
&\geq \sum_{i \in U} \mathbb{E}\left[\,\textsc{Mec}_t \mid L_i\,\right] p_i \cdot \alpha,
\end{aligned}$$

where the last inequality comes from the fact that the probability that every i satisfies $X_i < \tau_t$ is exactly α.

Suppose L_i happens. Then, we claim that under $\textsc{Mec}_t$ with τ_t, the agent's best-response is to propose a feasible set that contains an element exceeding τ_t. This follows from the fact that $\textsc{Mec}_t$ only accepts the set if there exists at least one element with $X_i \geq \tau_t$, and thus the agent's utility becomes positive only if he proposes a feasible set that contains i. Note that each singleton set $\{i\}$ is feasible due to the downward-closedness and the fact that we trim the set system at the beginning of the proof.

Further, in this case, the principal obtains the utility of X_i conditioned on E^c and L_i. This implies that $\mathbb{E}\left[\,\textsc{Mec}_t \mid L_i\,\right] = \mathbb{E}\left[\,X_i \mid L_i\,\right] = \mathbb{E}\left[\,X_i \mid X_i \geq \tau_t\,\right] = v_i$, due to the independence of X_i's. Hence we obtain $\mathbb{E}\left[\,\textsc{Mec}_t \mid E^c\,\right]\Pr\left[\,E^c\,\right] \geq \alpha \sum_{i \in U} p_i v_i$.

Therefore, we can decompose the principal's gain as follows.

$$\begin{aligned}\mathbb{E}\left[\,\textsc{Mec}_t\,\right] &= \mathbb{E}\left[\,\textsc{Mec}_t \mid E\,\right]\Pr\left[\,E\,\right] + \mathbb{E}\left[\,\textsc{Mec}_t \mid E^c\,\right]\Pr\left[\,E^c\,\right] \\ &\geq \mathbb{E}\left[\,\textsc{Mec}_t \mid E\,\right]\Pr\left[\,E\,\right] + \alpha \sum_{i\in[n]} p_i v_i. \end{aligned} \tag{A.3}$$

To upper bound $\textsc{Opt}$, we first observe that

$$\mathbb{E}\left[\sum_{i\in I^*\cap Z} X_i\right] \leq \mathbb{E}\left[\,X_{I^*}|E\,\right]. \tag{A.4}$$

This can be seen by, given arbitrary realization $(X_i)_{i\in U}$, we re-sample X_i for $i \in Z^c$ until they belong to Z. Such procedure will only increase $\mathbb{E}\left[\sum_{i\in I^*\cap Z} X_i\right]$, but this eventually computes $\mathbb{E}\left[\,X_{I^*}|E\,\right]$.

Hence, we obtain the following inequalities

$$\begin{aligned}\mathbb{E}\left[\,\textsc{Opt}\,\right] &= \mathbb{E}\left[\sum_{i\in I^*} X_i\right] \\ &= \mathbb{E}\left[\sum_{i\in I^*\cap Z} X_i\right] + \mathbb{E}\left[\sum_{i\in I^*\cap Z^c} X_i\right] \\ &\leq \mathbb{E}\left[\sum_{i\in I^*\cap Z} X_i\right] + \mathbb{E}\left[\sum_{i\in Z^c} X_i\right]. \end{aligned} \tag{A.5}$$

Meanwhile, observe that $\mathbb{E}\left[\sum_{i\in Z^c} X_i\right]$ is the expectation of the random variables that satisfy $X_i \geq \tau_t$, *i.e.,* $\mathbb{E}\left[\,X_i \mid X_i \geq \tau_t\,\right]$. Hence, we can further upper bound (A.5) by

$$\begin{aligned}\mathbb{E}\left[\,\textsc{Opt}\,\right] &\leq \mathbb{E}\left[\sum_{i\in I^*\cap Z} X_i\right] + \sum_{i\in U} p_i v_i \\ &\leq \mathbb{E}\left[\,X_{I^*} \mid E\,\right] + \sum_{i\in U} p_i v_i. \qquad \text{(By (A.4))} \end{aligned}$$

Assume $\mathbb{E}\left[\, X_{I^*} \mid E \,\right] \leq \sum_{i \in [n]} p_i v_i$. Then, we simply use $\textsc{Mec}_t$ with τ_t. This gives us $\alpha/2$-approximation since (A.3) guarantees that $\textsc{Mec}_t$ obtains at least $\alpha \sum_{i \in U} p_i v_i$, whereas $\mathbb{E}\left[\, \textsc{Opt} \,\right] \leq 2 \sum_{i \in U} p_i v_i$. Given a reasonable choice of constant α, this gives a constant factor approximation.

Thus, we move on to the next case, and assume we have $\mathbb{E}\left[\, X_{I^*} \mid E \,\right] > \sum_{i \in [n]} p_i v_i$. Similarly, it suffices to bound lower bound a constant fraction of $\mathbb{E}\left[\, X_{I^*} \mid E \,\right]$. We further define some notations. Throughout the proof, we abuse $A|H$ to denote the random variable A conditioned on the event H. Let A_I be the characteristic vector of the feasible set I, *e.g.,* if $I = \{1\}$, then $A_I = (1, 0, 0, \ldots, 0)$. Then, given the random vector $X = (X_1, X_2, \ldots, X_n)|E$ conditioned on E, the optimal basis $X_{I^*}|E$ can be computed as

$$X_{I^*}|E = \max_{I \in \mathcal{I}} A_I \cdot X|E,$$

which is a pointwise maximum of the weights A_I. Conditioned on the event E, we know that every random variable X_i is bounded above by τ_t.

Normalizing the domain of our random variables $X_i|E$, define $Z_i = (X_i/\tau_t)|E$, *i.e.,* the random variables that is tailored to the event E, and normalized by τ_t. Further, define $Z_I = (X_I/\tau_t)|E$ for feasible set I, and $U_{-i} = U \setminus \{i\}$. To apply the concentration by Lemma A.4.1, we can set (i) $Z = \max_{I \in \mathcal{I}} Z_I$, and $Z^{(i)} = \max_{I \in \mathcal{I}|U_{-i}} Z_I$. Obviously, Z satisfies 1-Lipschitzness.

Let us see why Z satisfies self-bounding property. Fix any realization of $(Z_i)_{i \in U}$ and let I^* be the corresponding maximum weight feasible set. Define $Z_{-i} = (Z_1, Z_2, \ldots, Z_{i-1}, 0, Z_{i+1}, \ldots, Z_n)$

and $Z_0 = (Z_1, \ldots, Z_n)$.

$$\begin{aligned}\sum_{i \in U} \frac{1}{n-1} Z^{(i)} &= \sum_{i \in U} \frac{1}{n-1} \max_{I \in \mathcal{I} | U_{-i}} \sum_{i \in I} Z_i \\ &\geq \sum_{i \in U} \frac{1}{n-1} A_{I^*} \cdot Z_{-i} \\ &= A_{I^*} \sum_{i \in U} \frac{1}{n-1} Z_{-i} = A_{I^*} Z_0 = Z.\end{aligned}$$

Therefore we have $\sum_{i \in U} Z - Z^{(i)} \leq Z$, which implies that Z is self-bounding.

Thus, we can apply the concentration given by Lemma A.4.1 as follows:

$$\begin{aligned}\Pr\left[X_{I^*} \leq (1-\delta)\mathbb{E}\left[X_{I^*} \mid E \right] \mid E \right] &= \Pr\left[\tau_t \cdot Z_{I^*} \leq (1-\delta)\tau_t \cdot Z_{I^*} \right] \\ &= \Pr\left[Z_{I^*} \leq (1-\delta) Z_{I^*} \right] \\ &\leq \exp\left(-\frac{\delta^2 \mathbb{E}\left[Z_{I^*} \right]}{2} \right) \\ &= \exp\left(-\frac{\delta^2 \mathbb{E}\left[X_{I^*} \mid E \right]}{2\tau_t} \right),\end{aligned}$$

therefore we have

$$\mathbb{E}\left[X_{I^*} \geq (1-\delta)\mathbb{E}\left[X_{I^*} \mid E \right] \mid E \right] \geq 1 - \exp\left(-\frac{\delta^2 \mathbb{E}\left[X_{I^*} \mid E \right]}{2\tau_t} \right) \tag{A.6}$$

Assume $\tau_t \geq \beta \cdot \mathbb{E}\left[X_{I^*} \mid E \right]$ for some $\beta \geq 0$. Recall that we assume $\mathbb{E}\left[X_{I^*} \mid E \right] \geq \sum_{i \in U} p_i v_i$, which implies that $\mathbb{E}\left[X_{I^*} \mid E \right] \geq \mathbb{E}\left[\textsc{Opt} \right]/2$. Suppose that we are running $\textsc{Mec}_t$ with τ_t again. In

this case, our threshold τ_t is sufficiently large so that

$$\mathbb{E}\left[\,\textsc{Mec}_t\,\right] \geq \Pr\left[\,\text{at least one } i \text{ satisfies } X_i \geq \tau_t\,\right] \cdot \tau_t \geq \tau_t \cdot \alpha$$
$$\geq \frac{\alpha\beta}{2} \cdot \mathbb{E}\left[\,\textsc{Opt}\,\right].$$

This implies that $\textsc{Mec}_t$ is $\alpha\beta/2$-approximation.

Thus, we focus on the final case such that $\tau < \beta \cdot \mathbb{E}\left[\,X_{I^*}\,\right]$. From (A.6), we obtain

$$\mathbb{E}\left[\,X_{I^*} \geq (1-\delta)\mathbb{E}\left[\,X_{I^*} \mid E\,\right] \mid E\,\right] \geq 1 - \exp\left(-\frac{\delta^2}{2\beta}\right).$$

This implies that $\max_{I\in\mathcal{I}} X_I|E$ is at least $1-\delta$ fraction of its expectation with probability denoted in the RHS. Hence, running $\textsc{Mec}_c$ with $\tau_c = (1-\delta)\mathbb{E}\left[\,X_{I^*} \mid E\,\right]$ yields $(1-\delta)\cdot(1-e^{-\delta^2/2\beta})$-approximation.

Overall, our mechanism can be depicted as follows.

1. If $\mathbb{E}\left[\,X_{I^*} \mid E\,\right] \leq \sum_{i\in U} p_i v_i$, then we run $\textsc{Mec}_t$ with τ_t.
2. If $\mathbb{E}\left[\,X_{I^*} \mid E\,\right] > \sum_{i\in U} p_i v_i$, if $\tau \geq \beta \cdot \mathbb{E}\left[\,X_{I^*} \mid E\,\right]$, we still run $\textsc{Mec}_t$ with τ_t.
3. Otherwise, we run $\textsc{Mec}_c$ with $\tau_c = (1-\delta)\mathbb{E}\left[\,X_{I^*} \mid E\,\right]$.

Roughly optimizing over the choices of α, β and δ, we obtain $1/9.1$-approximation. a

A.4.2 Proof of Theorem 1.6.1

Proof of Theorem 1.6.1. Let $\textsc{Alg}$ be an α-competitive non-adaptive algorithm for the matroid prophet inequality. For ease of exposition, let us first assume that $\textsc{Alg}$ is deterministic, and in the end of the proof we will discuss the extension to randomized algorithms. For simplicity, we denote U by

$[n] = \{1, 2, \ldots, n\}$. Consider the following procedure to construct the mechanism.

Construction of the mechanism

i) Given $(X_i, Y_i)_{i\in[n]}$, construct an instance of matroid prophet inequality such that $W_i \sim X_i$ and $T_i \sim e^{-Y_i}$ for $i \in U$.

ii) Given the constructed instance, call ALG on $(\mathcal{M}, W, T)$ to obtain the non-adaptive thresholds τ_i for $i \in [n]$.

iii) Define the eligible set as $\mathcal{E} = \{(I, X, Y) : X_i \geq \tau_i, \forall i \in I, I \in \mathcal{I}\}$.

Let us denote the constructed mechanism by MEC. Note that MEC is well-defined since the thresholds τ_i are fixed before the random variables are realized. Furthermore, obtaining τ_i only requires knowledge about X_i since ALG is order-oblivious. Hence, constructing the eligible set $\mathcal{E}$ does not require any information regarding the agent's utility Y_i nor its realization, *i.e.,* MEC is a well-defined agent-oblivious mechanism.

Now we prove that MEC is α-approximate against OPT. Given any realization of the random variables $(X_i, Y_i)_{i\in[n]}$, let the resulting allocation of ALG for the matroid prophet inequality be $A \in \mathcal{I}$. Due to the nature of ALG, A consists of elements satisfying $w_i \geq \tau_i$ with respect to the arrival order $T_i \sim e^{-Y_i}$ and the matroid constraint. In the delegation procedure with MEC, denote the agent's best-response given the realizations $(X_i, Y_i)_{i\in[n]}$ by $A^* \in \mathcal{I}$. By definition, it maximizes the agent's utility, *i.e.,* $A^* = \operatorname{argmax}_{I\in\mathcal{I}, X_i\geq\tau_i, \forall i\in I} \sum_{i\in I} Y_i$.

Now we prove that the best-response A^* is exactly the same as the allocation A. Fix a realization of the utility random variable for each solution. Under the constraint on the eligible set $\mathcal{E}$, let V be the

set of eligible elements. By Lemma A.1.1, the set-system $\mathcal{M}|V$ restricted to V is also a matroid. Let $\mathcal{M}' = \mathcal{M}|V = (V, \mathcal{I}')$ where $\mathcal{I}' = \mathcal{I}(\mathcal{M}|V)$.

We proceed by proving that the agent's best response A^* coincides with the output S of the greedy algorithm 4 with weights Y_i for each $i \in [n]$. Recall that the agent's best-response is an independent set I that maximizes the agent's utility, *i.e.,* $A^* = \operatorname{argmax}_{I \in \mathcal{I}, X_i \geq \tau_i, \forall i \in I} \sum_{i \in I} Y_i$. By Lemma A.1.2, the independent set S obtained by Algorithm 4 with input matroid $\mathcal{M}'$ satisfies $S = \operatorname{argmax}_{I \in \mathcal{I}'} \sum_{i \in I} Y_i$. Note that this procedure is well-defined since $\mathcal{M}|V$ is also a matroid. By our construction of $\mathcal{M}|V$, the constraint $\{I : I \in \mathcal{I}, X_i \geq \tau_i, \forall i \in I\}$ is the same as $\{I : I \in \mathcal{I}'\}$, which implies that $A^* = S$.

We further prove that the output of the greedy algorithm under matroid $\mathcal{M}'$ is the same as the independent set obtained by the algorithm ALG in the prophet inequality. If A contains the first element 1, this implies that $X_1 \geq \tau_1$, and thus $1 \in \mathcal{M}'$. Therefore, 1 is included in S too. If A does not contain 1, then it must be the case that $X_1 < \tau_1$. Then, $1 \notin V$ and thus S does not contain 1 as well. To prove by induction, for some $i \in 2, \ldots, n$, suppose that the outcome of ALG in the prophet inequality and of the greedy algorithm with $\mathcal{M}'$ are the same up to element $i - 1$. Note that the base case is shown by the above argument. If A contains element i, similarly, we have $X_i \geq \tau_i$, thereby implying $i \in S$. Otherwise if $i \notin A$, then it must satisfy $X_i < \tau_i$ and thus $i \notin S$. Hence A and S must be the same up to element i as well, and this concludes the proof.

Combining two claims above, we observe that $A = A^*$ since they both coincide with the output S of the greedy algorithm under $\mathcal{M}'$.

If ALG is a randomized algorithm, the obtained non-adaptive thresholds τ_i will be randomized, *i.e.,* for each $i \in [n]$ step ii) in the mechanism description will return the thresholds $\tau^{(j)} \in \mathbb{R}^n_{\geq 0}$ and corresponding probability $p_j \in [0, 1]$. If the random bits of ALG and MEC are coupled in the coupled

problem instances, one can see that the final outcome will be exactly the same given any realization of the random bits. This is because given the coupled random bits, the resulting thresholds in ALG will be equivalent to the thresholds in MEC, so we can apply the deterministic version of the argument shown above. Since this argument can be applied for any realization of the random bits, the outcomes of ALG and MEC will be equivalent. Thus in expectation, they obtain the same approximation ratio, and it finishes the proof. a

A.4.3 Proof of Corollary 1.6.3

Number of units	1	2	3	4	5	6
Competitive ratio	0.5	0.585	0.630	0.660	0.682	0.698

Table A.1: Competitive ratio for small k in Corollary 1.6.3 and Lemma A.4.2, obtained by [31].

The proof of Corollary 1.6.3 requires the following lemma.

Lemma A.4.2 ([33] and [31]). *There exists a static algorithm for k-unit prophet inequality with competitive ratio of $1 - O(\sqrt{\frac{\log k}{k}})$. In particular, let X be a Poisson random variable with a rate defined such that the following equation holds:*

$$\frac{1}{k}\mathbb{E}\left[\min(X, k)\right] = \Pr\left[X \leq k - 1\right].$$

Then, the competitive ratio is either side of the equation. For small k, the resulting values are shown in Table A.1.

Combining Theorem 1.6.1 and lemma A.4.2 immediately implies Corollary 1.6.3.

A.4.4 Proof of Theorem 1.7.2

To prove the theorem, we first need the following variant of Lemma A.1.2, that adapts to the setting where there exists a threshold for each element which may prevent an algorithm picking the element.

Lemma A.4.3. *Let $\mathcal{I}$ be a collection of subsets of U. Given thresholds $\tau = (\tau_1, \ldots, \tau_{|U|}) \in \mathbb{R}^{|U|}_{\geq 0}$, $(U, \mathcal{I})$ is a matroid if and only if the following properties hold:*

1. *$\emptyset \in \mathcal{I}$.*
2. *If $I \in \mathcal{I}$ and $I' \subset I$ then $I' \in \mathcal{I}$.*
3. *For all weight functions $w : U \mapsto \mathbb{Z}_{\geq 0}$, running the greedy algorithm 4 in the restricted matroid $\mathcal{M}|V$ where $V = \{e \in U : w_e \geq \tau_e\}$ with weight w produces a maximal member of $\mathcal{I}$ with maximum weight among $\mathcal{M}|V$.*

Proof of the lemma. If $\tau = (0, 0, \ldots, 0)$, then the lemma statement is essentially equivalent to Lemma A.1.2, which finishes the proof. Hence, we assume that at least one coordinate in τ is nonzero.

Also, the only if part follows from the property 3 in Lemma A.1.2 since restricted matroid $\mathcal{M}|V$ is always a matroid as well, so we focus on the other side in the proof.

To show that $\mathcal{M} = (U, \mathcal{I})$ is a matroid if all the three properties hold, we prove by contradiction. Assume that the augmentation property does not hold for $I_1, I_2 \in \mathcal{I}$, *i.e.,* for some $I_1, I_2 \in \mathcal{I}$ we have $|I_1| < |I_2|$ and $I_1 \cup e \notin \mathcal{I}$ for every $e \in I_2 - I_1$. We write $w(I)$ to denote the weight of $I \in \mathcal{I}$, *i.e.,* $w(I) = \sum_{e \in I} w_e$. Let $\tau^* = \max_{e \in U} \tau_e$, which is a positive rational number as we assume that at least one coordinate in τ is strictly positive. Since $|I_1 - I_2| < |I_2 - I_1|$, we can choose a rational

number $\varepsilon > 0$ such that $\varepsilon = q/p$ with $p, q \geq \tau^*$ and

$$\frac{|I_1 - I_2|}{|I_2 - I_1|} < \varepsilon < 1. \tag{A.7}$$

Let $\varepsilon = q/p$ for positive integers $p, q \in \mathbb{Z}_{>0}$, where $q < p$ and p and q are coprime. Then, we have

$$p > q \geq \tau^* \geq 1. \tag{A.8}$$

$$p \cdot \varepsilon = q \tag{A.9}$$

Now, define the weight function $w : U \mapsto \mathbb{R}$ by

$$w_e = \begin{cases} p & \text{if } e \in I_1 \\ q & \text{if } e \in I_2 - I_1 \\ 0 & \text{otherwise.} \end{cases}$$

Due to (A.8), the output of the greedy algorithm in this case contains I_1, since each element $e \in I_1$ satisfies $w_e \geq \tau^* \geq \tau_e$, $I_1 \in \mathcal{I}$, and they have the largest weights. Further, by assumption, it cannot select any more element with positive weights in $I_2 - I_1$. Thus, the greedy exactly selects I_1. By the downward-closedness of $\mathcal{M}$, however, I_2 will be contained in a maximal feasible set $I_2' \in \mathcal{I}$, and we

have

$$\begin{aligned} w(I_2') \geq w(I_2) &= |I_1 \cap I_2| p + |I_2 - I_1| q \\ &= |I_1 \cap I_2| p + |I_2 - I_1| \cdot q \cdot \varepsilon && \text{(By (A.9))} \\ &> |I_1 \cap I_2| p + |I_2 - I_1| \cdot q \cdot \frac{|I_1 - I_2|}{|I_2 - I_1|} = w(I_1). && \text{(By (A.7))} \end{aligned}$$

Note that the last equation follows from

$$w(I_1) = p \cdot |I_1| = p \cdot (|I_1 \cap I_2| + |I_1 - I_1|).$$

Therefore, I_2' is the maximal member with maximum weights, contradicting the third property. Thus, the augmentation property should hold, which concludes that $\mathcal{M}$ is a matroid. a

Now we can prove Theorem 1.7.2 as follows.

Proof of Theorem 1.7.2. Note that the first and second property hold as $\mathcal{I}$ is downward-closed. Thus, it suffices to prove that if ALG is implementable, then the greedy algorithm outputs the maximum weight member of $\mathcal{I}$ for any non-negative integer weights by Lemma A.4.3. Since ALG is implementable, for every problem instance $(\mathcal{M}, X, Y)$, the following should hold

$$I^* := A(\textsc{Alg}; \mathcal{M}, W, T) = B(\textsc{Mec}; \mathcal{M}, X, Y), \tag{A.10}$$

where MEC is the induced mechanism from ALG, under the coupling such that $W_i \sim X_i$ and $T_i \sim e^{-Y_i}$ for $i \in U$. Recall the definition of interim outcome as follows: given a realization (x, y) of random vectors (X, Y), an interim outcome of MEC is the outcome of the mechanism given the realization,

denoted by $g(\textsc{Mec}; \mathcal{M}, x, y) \in \mathcal{I} \times \left(\bigtimes_{i \in U}(\Phi_i, \Psi_i)\right)$. Similarly, we define the interim outcome of ALG for prophet inequality to be the outcome that denotes the allocated set and utility given the realization of the random variables. Denote this by $h(\textsc{Alg}; \mathcal{M}, w, t)$, where (w, t) denotes the realization of (W, T).

Then, (A.10) implies that

$$h(\textsc{Alg}; \mathcal{M}, w, t) = g(\textsc{Mec}; \mathcal{M}, x, y), \tag{A.11}$$

for any vectors w, t, x, y such that $w_i = x_i$ and $t_i = e^{-y_i}$ for $i \in U$ due to the coupling.

We consider a problem instance with discrete support whose probability exponentially decays. More precisely, define the support $S = \mathbb{Z}_{\geq 0}$ and the probability density function $f(\cdot)$ by $f(i) = 2^{-i-1}$ for every $i \in \mathbb{N}$. This is well-defined since $\sum_{n \in \mathbb{Z}_{\geq 0}} f(n) = 1$. Consider a problem instance where the random variables X_i and Y_i for every $i \in U$ follow this distribution. Obviously, the support of X_i and Y_i is the entire set of non-negative integers $\mathbb{Z}_{\geq 0}$. Let τ be the vector of thresholds obtained by running ALG on this problem instance. Let $\tau^* = \max_{e \in U} \tau_e$. Then, the realization x satisfies $x_e \geq \tau^*$ for every $e \in U$ with strictly positive probability, due to our construction of the density function $f(\cdot)$. Given such realization of $x = w$, note that every element $e \in U$ satisfies $w_e = x_e \geq \tau_e$. Further, due to the independence of X and Y, given the realization x satisfying $x_e \geq \tau^*$ for every $i \in U$, every non-negative integer vectors $y \in \mathbb{Z}_{\geq 0}^{|U|}$ can be realized with strictly positive probability for Y.

Thus, conditioned on such x, for any $y \in \mathbb{Z}_{\geq 0}^{|U|}$, we should have the following equation by (A.11).

$$h(\textsc{Alg}; \mathcal{M}, w, t) = g(\textsc{Mec}; \mathcal{M}, x, y), \tag{A.12}$$

given $w_i = x_i$ and $t_i = e^{-y_i}$ for every $i \in U$.

Since every element satisfies $w_e \geq \tau_e$, the element e for ALG arrive in the decreasing order of y_e. Thus, ALG selects the elements in U in the decreasing order of y_e.

Meanwhile, the induced mechanism MEC also imposes the same thresholds τ on x, so every element satisfies the threshold constraint $x_e \geq \tau_e$. Note that the agent's best-response is to select a maximal feasible set $I \in \mathcal{I}$ with maximum weights w.r.t. the weights y. That is, $g(\text{MEC}; \mathcal{M}, x, y)$ is the maximal feasible set in $\mathcal{M}$ with maximum weights with respect to y. This implies that running the greedy algorithm in U outputs the maximal feasible set with maximum weight, for arbitrarily selected non-negative integer weights y. This satisfies the property 3 of Lemma A.4.3, thereby concluding that $\mathcal{M}$ is a matroid.

a

A.4.5 Proof of Proposition 1.8.1 and Corollary 1.8.2 and 1.8.3

Proof of Proposition 1.8.1. Recall that a matroid-constrained non-adaptive algorithm imposes a sub-constraint on the feasible set upon the non-adaptive thresholds, where the constraint itself forms a matroid. Mainly, we can consider the following construction of a mechanism MEC, given an oracle access to a matroid-constrained non-adaptive order-oblivious algorithm ALG. Suppose ALG is deterministic.

Construction of the mechanism

i) Given $(X_i, Y_i)_{i\in[n]}$, construct an instance of matroid prophet inequality such that $W_i \sim X_i$ and $T_i \sim e^{-Y_i}$.

ii) Given the constructed instance, call ALG on (W, σ) to obtain the following:

   (a) Non-adaptive thresholds τ_i for $i \in [n]$.

   (b) Sub-constraint $\mathcal{J} \subseteq \mathcal{I}$ such that $\mathcal{J}$ is also a matroid.

iii) Define eligible set $\mathcal{E} = \{(J, X, Y) : X_j \geq \tau_j, \forall j \in J, J \in \mathcal{J}\}$.

Note that MEC is well-defined to be agent-oblivious mechanism since ALG only requires the knowledge on the distribution of $W_i \sim X_i$ for $i \in U$.

Analogously to the proof of Theorem 1.6.1, it suffices to prove that the resulting collection of eligible sets constitutes a matroid. This can be seen by first observing that $\mathcal{J}$ is a matroid, then given the non-adaptive thresholds $(\tau_i)_{i\in[n]}$, we can collect a set V consisting of feasible elements i satisfying $X_i \geq \tau_i$, and finally the restriction $\mathcal{J}|V$ is still a matroid by Lemma A.1.1. The argument for randomized algorithm easily carries over, too. a

Proof of Corollary 1.8.2. The proof directly follows from the proof of Theorem 12 in [34]. Mainly, they partition the corresponding edges of the graphic matroid, and view each partition as a 1-uniform matroid so that the algorithm picks only at most one pre-specified edge from each partition. Thus, the set of feasible edges essentially constitutes a restricted set system over the original matroid, which itself is a matroid by Lemma A.1.1. a

Proof of Corollary 1.8.3. We briefly recap how the OCRS partitions the matroid into matroid minors.

Formally, given a matroid $\mathcal{M} = (U, \mathcal{I})$, let $P_{\mathcal{M}}$ be the associated matroid polytope. Then, we randomly sample a subset J from U, and designate the elements in J as active elements.[2] We limit our attention to the restricted matroid $\mathcal{M}' = \mathcal{M}|J$.

Mainly, [52] construct a chain decomposition of the elements.

$$\emptyset = U_\ell \subsetneq U_{\ell-1} \subsetneq \ldots \subsetneq U_1 \subsetneq U_0 = U.$$

Then, they accepts an active element $e \in U_i \setminus U_{i+1}$, if given the already accepted elements in $U_i \setminus U_{i+1}$, adding e to the current set forms an independent set in the matroid obtained from contracting $\mathcal{M}$ by U_{i+1} and then restricted to N_i, *i.e.,* $(\mathcal{M}/U_{i+1})|N_i$. Alternatively, this can be viewed as defining a collection of feasible sets as

$$\mathcal{C} = \{I \in \mathcal{I} : \text{for every } i \in \{0, 1, \ldots, \ell-1\}, I \cap (U_i \setminus U_{i+1}) \in \mathcal{I}((\mathcal{M}'/U_{i+1})|U_i)\}.$$

It suffices to prove that $\mathcal{C}$ is a matroid. Equivalently, one can define $\mathcal{C}$ as the following.

$$\begin{aligned}\mathcal{C} &= \{\cup_{i=0}^{\ell-1} I_i \in \mathcal{I} : I_i \in \mathcal{I}(\mathcal{M}'/U_{i+1}|U_i), I_i \subset U_i \setminus U_{i+1}\} \\ &= \{\cup_{i=0}^{\ell-1} I_i \in \mathcal{I} : I_i \in \mathcal{I}(\mathcal{M}'/U_{i+1}|U_i)\},\end{aligned}$$

where the second equation follows from the fact that the domain of $\mathcal{M}'/U_{i+1}|U_i$ is in fact equivalent to $U_i \setminus U_{i+1}$. Note that $(\mathcal{M}'/U_{i+1})|U_i$ is a matroid by Lemma A.1.1. Let $\mathcal{M}_i = (\mathcal{M}'/U_{i+1})|U_i$ for each $i = 0, 1, \ldots, \ell-1$. Then, $\mathcal{C}$ is indeed the union of matroids, *i.e.,* $\mathcal{C} = \cup_{i=0}^{\ell-1} \mathcal{M}_i$. Thus by Lemma A.1.3,

[2]The notion of active element for the reduction between OCRS and matroid prophet inequality can be found in Section 4.1 in [52]. We also present the details in the proof of Theorem **??**, when we discuss the direct connection to OCRS to obtain constant approximate mechanisms beyond matroid.

we conclude that $\mathcal{C}$ is a matroid. a

A.4.6 Proof of Theorem 1.8.4

To prove Theorem 1.8.4, we first require the following technical lemma.

Lemma A.4.4. *Suppose there exists an agent-oblivious (-aware) mechanism* MEC_c *which is* α*-approximate for problem instances such that the* X_i *is equipped with an absolutely continuous distribution with density function bounded above. Then, for any* $\delta > 0$*, there exists an* $(\alpha - \delta)$*-agent-oblivious (-aware) approximate mechanism* MEC_g *for the delegated choice under the same constraint* $\mathcal{I}$*.*

Proof of Lemma A.4.4. The main idea behind this result is to add a small noise to the input distribution to make it continuous. Then, we construct a mechanism that simulates the given mechanism under the perturbed distribution, and imposes the eligible set based on its outcome. In particular, since the realizations of the random variables are only observed by the misaligned agent while it is the principal who designs the mechanism in advance, the technical subtlety lies on analyzing the agent's best response given the perturbed distribution.

Precisely, consider a problem instance $(\mathcal{M}, X, Y)$ of the delegated choice with set system $\mathcal{M} = (U, \mathcal{I})$, where $n = |U|$. Let $D_1, \ldots, D_n$ denote the distributions of the agent's utilities $X_1, \ldots, X_n$. Consider a Gaussian noise vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ where each ε_i is drawn from $\mathcal{N}(0, \sigma)$, where $\sigma > 0$ will be specified later. Define $X'_i = X_i + |\,\varepsilon_i\,|$ and let D'_i denote the distribution of X'_i for $i \in U$. By Lemma A.3.4, D'_i is absolutely continuous for $i \in U$. Consider running MEC_c under the problem instance $(\mathcal{M}, X', Y)$, *i.e.,* the principal's utility distributions are absolutely continuous. Without loss of generality, assume that MEC_c is single-proposal mechanism by Proposition 1.4.2, and let $\mathcal{E}_c$ be the eligible set of MEC_c.[3]

[3] Alternatively, one can adapt the procedure illustrated in the proof of Proposition 1.4.2 to convert the mechanism to the

Denote the problem instance $(\mathcal{M}, X, Y)$ by Θ_g and that of $(\mathcal{M}, X', Y)$ by Θ_c. We further write $x(I, \Theta)$ to denote the principal's utility of set $I \in \mathcal{I}$ under the problem instance Θ, and similarly $y(I, \Theta)$ for the agent's utility.

Let I_c denote the outcome of $\textsc{Mec}_c$ on the problem instance Θ_c. Since $\textsc{Mec}_c$ is α-approximate on Θ_c, we have

$$\frac{\mathbb{E}\left[\, x(I_c, \Theta_c) \,\right]}{\mathbb{E}\left[\, \max_{I \in \mathcal{I}} x(I, \Theta_c) \,\right]} \geq \alpha.$$

We design the mechanism $\textsc{Mec}_g$ as follows.

- The mechanism first independently samples ε_i from $\mathcal{N}(0, \sigma)$ for each $i \in U$.
- It announces the eligible set $\mathcal{E}_g = \{(I, X, Y) : \mathcal{E}_c(I, X', Y) = 1\}$.

Also, the construction of the mechanism implies that if $\textsc{Mec}_c$ is agent-oblivious then $\textsc{Mec}_g$ is agent-oblivious and if $\textsc{Mec}_c$ is agent-aware then $\textsc{Mec}_g$ should be also made agent-aware.

First, we have

$$\mathbb{E}\left[\max_I x(I, \Theta_g) \right] \leq \mathbb{E}\left[\max_I x(I, \Theta_c) \right] + \mathbb{E}\left[\sum_{i \in U} |\, \varepsilon_i \,| \right] \leq \mathbb{E}\left[\max_I x(I, \Theta_c) \right] + O(n\sigma).$$

Recall that the agent's best response is to propose a set that maximizes his own expected utility, *i.e.,*

$$I_g := \operatorname*{argmax}_{I \in \mathcal{I} : \mathcal{E}_g(I, X, Y) = 1} y(I, \Theta_g).$$

single-proposal mechanism.

Due to our construction of MEC_g, observe that

$$\begin{aligned} I_g = \operatorname*{argmax}_{I \in \mathcal{I}: \mathcal{E}_g(I,X,Y)=1} y(I, \Theta_g) &= \operatorname*{argmax}_{I \in \mathcal{I}: \mathcal{E}_c(I,X',Y)=1} y(I, \Theta_g) \\ &= \operatorname*{argmax}_{I \in \mathcal{I}: \mathcal{E}_c(I,X',Y)=1} y(I, \Theta_c), \end{aligned}$$

where the second equality comes from the fact that Θ_c and Θ_g are indifferent with respect to the agent's utility. Hence, I_g under Θ_g is equivalent to I_c under Θ_c. This implies that

$$\begin{aligned} x(I_g, \Theta_g) = x(I_c, \Theta_c) - \sum_{i \in I_c} |\varepsilon_i| &= \sum_{i \in I_c} X_i' - \sum_{i \in I_c} |\varepsilon_i| \\ &\geq \sum_{i \in I_c} X_i' - \left(\sum_{i \in [n]} |\, \varepsilon_i \,| \right) \end{aligned}$$

Thus, we obtain

$$\begin{aligned} \mathbb{E}\,[\, x(I_g, \Theta_g)\,] &\geq \mathbb{E}\,[\, x(I_c, \Theta_c)\,] - \mathbb{E}\left[\sum_{i \in [n]} |\, \varepsilon_i \,| \right] \\ &\geq \mathbb{E}\,[\, x(I_c, \Theta_c)\,] - O(n\sigma). \end{aligned}$$

Therefore, we have

$$\frac{\mathbb{E}\,[\, x(I_g, \Theta_g)\,]}{\mathbb{E}\,[\, \max_I x(I, \Theta_g)\,]} \geq \frac{\mathbb{E}\,[\, x(I_c, \Theta_c)\,] - O(n\sigma)}{\mathbb{E}\,[\, \max_I x(I, \Theta_c)\,] + O(n\sigma)},$$

which, letting $\sigma \mapsto 0$ tends to

$$\frac{\mathbb{E}\left[\, x(I_c, \Theta_c) \,\right]}{\mathbb{E}\left[\, \max_I x(I, \Theta_c) \,\right]} \geq \alpha.$$

Thus, for any $\delta > 0$, we can pick small enough $\sigma > 0$ to obtain the desired bound. a

Now we are ready to prove Theorem 1.8.4

Proof of Theorem 1.8.4. The overall argument to construct such mechanism can be depicted as follows.

Construction of the mechanism

i) Let ALG be the α-competitive boundary-randomized order-oblivious non-adaptive algorithm for matroid prophet inequality.

ii) Given a problem instance $(\mathcal{M}, X, Y)$ of the matroid delegated choice, consider another problem instance $(\mathcal{M}, X', Y)$ such that $X'_i \sim X_i + |\,\varepsilon_i\,|$, where ε_i is Gaussian noise defined in the proof of Lemma A.4.4.

iii) Construct a problem instance $(\mathcal{M}, W', T)$ of matroid prophet inequality such that $W'_i \sim X'_i$ and $T_i \sim e^{-Y_i}$.

iv) Since ALG behaves exactly as non-adaptive algorithm under the problem instance $(\mathcal{M}, W', T)$ with continuous distributions for W', we call ALG on $(\mathcal{M}, W', T)$ to obtain non-adaptive thresholds τ.

v) Construct the intermediate mechanism MEC$'$ as follows.

   a) Announce thresholds τ.

   b) Accept the proposed I if $I \in \mathcal{I}$ and $X_i \geq \tau_i$ for every $i \in U$.

vi) Apply Lemma A.4.4 with MEC$'$ to obtain MEC.

By the construction of MEC$'$, it is α-approximation agent-oblivious matroid delegation mechanism under the problem instances $(\mathcal{M}, X', Y)$ such that the principal's utility random variables have continuous probability distributions. Thus, MEC can be constructed as a $(\alpha - \varepsilon)$-approximate agent-oblivious mechanism, for $\varepsilon > 0$. a

A.4.7 Proof of Proposition A.2.1 and Corollary A.2.2 and A.2.3

Proof of Proposition A.2.1. Suppose ALG is deterministic. The randomized algorithm case will be handled analogously to the proof of Theorem 1.6.1, and we omit the details. Consider the following procedure to construct the mechanism.

Construction of the mechanism

i) Given $(\mathcal{M}, X, Y)$ of delegated choice, construct an instance of matroid prophet such that $W_i \sim X_i$ and $T_i \sim e^{-Y_i}$, for each $i \in [n]$.

ii) Given the constructed instance $(\mathcal{M}, W, T)$, call ALG on the matroid prophet inequality to obtain the thresholds τ.

iii) Define eligible set $\mathcal{E} = \{(I, X, Y) : X_i \geq \tau_i, \forall i \in I, I \in \mathcal{I}\}$.

Note that MEC is well-defined agent-aware mechanism since MEC's knowledge on Y implies the information for the distribution of T, thus we can call ALG with this informatoin on the distribution of T.

Analogously to the proof of Theorem 1.6.1, it suffices to prove that the resulting collection of eligible sets constitutes a matroid. This can be seen by noting that given the non-adaptive thresholds $(\tau_i)_{i\in[n]}$, we can collect a set V that consists of feasible elements satisfying $X_i \geq \tau_i$, and finally the restriction $\mathcal{M}|V$ is still a matroid by Lemma A.1.1. The argument for a randomized algorithm easily carries over, and we skip the details. a

Proof of Corollary A.2.2. The proof follows from observing that if Y_i is deterministic and the principal

knows it, the step i) in the proof of Theorem A.2.1 induces a deterministic arrival order T, and thus we can use an order-aware algorithm to implement the mechanism. a

Proof of Corollary A.2.3. Since we consider an agent-aware mechanism, we have the information that $Y_i \sim Y_j$ for every $i, j \in U$. Thus, following the construction of the mechanism in the proof of Theorem A.2.1, step i) outputs an arrival order T_i such that all the T_i have the same distribution for $i \in U$. Given the arrival times T, let the induced arrival order be $\sigma \in P_n$ where P_n is the set of all permutations over $[n]$. Then, it is easy to see that the resulting distribution over σ is a uniform distribution over P_n. Thus, the resulting matroid prophet inequality problem is the matroid prophet secretary problem. The remaining arguments are analogous to the proof of Theorem A.2.1, and we omit the details. a

A.4.8 Proof of Theorem 1.9.1

Proof of Theorem 1.9.1. Overall, we consider a game between an algorithm designer and an adversary who chooses the orderings. An action by the designer is a deterministic non-adaptive algorithm while an action by the adversary is an ordering of the elements. The designer tries to maximize the expected output of the algorithm, where the expectation is over the randomness in the weight of the elements, while the adversary tries to minimize it. A randomized strategy by the agent is a distribution over the deterministic non-adaptive thresholds, while a randomized strategy by the adversary is a distribution over the arrival orders. If the agent moves first, then it needs to design an algorithm without knowing the arrival order, *i.e.,* order-oblivious algorithm. If the agent moves second, however, it designs the algorithm while knowing the distribution of the arrival order, *i.e.,* order-hinted algorithm.

More formally, let $\mathcal{A}$ be the collection of all possible non-adaptive algorithms for the prophet

inequality problem with set system $\mathcal{M}$, and let P_n be the set of all possible permutations defined over $[n]$. Write $\Delta(\mathcal{A})$ and $\Delta(P_n)$ to denote the probability simplex over $\mathcal{A}$ and P_n, respectively. That is, $A \in \Delta(\mathcal{A})$ and $T \in \Delta(P_n)$ denotes the randomized non-adaptive algorithm and randomized arrival order, respectively. Note that it suffices to restrict the domain of arrival times to be P_n, since what only matters for the prophet inequality is the relative arrival order between the random variables. Let us fix the random weights W and set system $\mathcal{M}$, and we write $x(A, T)$ to denote the algorithm A's expected gain under the arrival order T. If the agent moves first, the outcome of the game will be

$$\inf_{T \in \Delta(P_n)} \sup_{A \in \Delta(\mathcal{A})} \mathbb{E}\left[\, x(A, T) \,\right]. \tag{A.13}$$

On the other hand, if the adversary moves first, the outcome of the game will be

$$\sup_{A \in \Delta(\mathcal{A})} \inf_{T \in \Delta(P_n)} \mathbb{E}\left[\, x(A, T) \,\right]. \tag{A.14}$$

Thus, if we could apply Lemma A.3.2, this would imply (A.13) and (A.14) are the same, thereby concluding that there exists an order-oblivious algorithm with the same approximation factor. However, the subtlety lies in the fact that $\mathcal{A}$ is not finite. Thus, we will show that there exists a reasonable class of algorithms that approximate $\mathcal{A}$ well.

Given the set system $\mathcal{M}$, let us fix the problem instance $(\mathcal{M}, W, T)$ and let D_i be the distribution of W_i for $i \in U$. Thanks to Lemma A.3.3, it is without loss of generality to assume that $(D_i)_{i \in U}$ are absolutely continuous and their density functions are uniformly bounded above by some constants. Let γ be the upper bound on the density functions of the distributions $D_1, \ldots, D_n$. Given any non-adaptive algorithm A, for some $\theta \in \mathbb{R}_{\geq 0}$, we define A_θ to be the same algorithm but thresholds larger than θ are

replaced with θ. For some natural number $m \in \mathbb{N}$, let us further define $A_{\theta,m}$ to be the algorithm A_θ such that the thresholds are rounded down to the closest integer multiples of $1/m$. We define $\mathcal{A}_\theta$ to be the collection of all possible algorithms A_θ, and let $\mathcal{A}_{\theta,m}$ be that of all possible algorithms $A_{\theta,m}$.

Claim A.4.5. *Given a problem instance $(\mathcal{M}, W, T)$ and any $\varepsilon > 0$, there exists $\theta \in \mathbb{R}_{\geq 0}$ such that for any $T \in \Delta(P_n)$ and any non-adaptive algorithm $A \in \Delta(\mathcal{A})$, we have*

$$\mathbb{E}\left[\, x(A_\theta, T) \,\right] \geq \mathbb{E}\left[\, x(A, T) \,\right] - \varepsilon.$$

Note that the threshold θ is independent of A and T.

Proof. Consider a threshold θ such that $\Pr\left[\, \max W_i \geq \theta \,\right] < \varepsilon'$. Let $x_i(A, T)$ denote the contribution of element i to the algorithm. We will show that

$$x_i(A_\theta, T) \geq (1 - O(\varepsilon))\, x_i(A, T),$$

with a proper choice of ε'. Denote the element i's threshold under A by τ. Further, let $\tau_\theta = \min(\theta, \tau)$. Then we have

$$\begin{aligned} x_i(A, T) &= \mathbb{E}\left[\, W_i \cdot \mathbf{1}\left\{\, i \text{ is feasible under } A \text{ and } W_i \geq \tau \,\right\}\right] \\ &= \Pr\left[\, i \text{ is feasible under } A \,\right] \mathbb{E}\left[\, w_i \cdot \mathbf{1}\left\{\, W_i \geq \tau \,\right\}\right], \end{aligned}$$

where the second equation comes from the fact that the event that i is feasible to be selected is deter-

mined by other elements' values, which is independent from W_i. Since $\tau_\theta \leq \tau$, we have

$$\mathbb{E}\left[\, W_i \cdot \mathbf{1}\left\{\, W_i \geq \tau \,\right\}\right] \leq \mathbb{E}\left[\, W_i \cdot \mathbf{1}\left\{\, W_i \geq \tau_\theta \,\right\}\right].$$

In addition, with probability at least $1 - \varepsilon'$, the event $\max_{j \neq i} W_j < \theta$ occurs. Therefore, with probability at least $1 - \varepsilon'$, the algorithm A and A_θ have the same output until the element i arrives. This implies that

$$\Pr\left[\, i \text{ is feasible under } A_\theta \,\right] \geq \Pr\left[\, i \text{ is feasible under } A \,\right] - \varepsilon'.$$

Therefore, we have

$$\begin{aligned} x_i(A_\theta, T) &= \Pr\left[\, i \text{ is feasible under } A_\theta \,\right] \mathbb{E}\left[\, W_i \cdot \mathbf{1}\left\{\, W_i \geq \tau \,\right\}\right] \\ &\geq \left(\Pr\left[\, i \text{ is feasible under } A \,\right] - \varepsilon\right) \mathbb{E}\left[\, W_i \cdot \mathbf{1}\left\{\, W_i \geq \tau_\theta \,\right\}\right] \\ &\geq x_i(A, T) - \varepsilon' \cdot \mathbb{E}\left[\, W_i \mathbf{1}\left\{\, W_i \geq \tau_\theta \,\right\}\right]. \end{aligned}$$

Plugging $\varepsilon' = \frac{\varepsilon}{\mathbb{E}[\max_{i \in U} W_i]}$, we obtain

$$\varepsilon' \cdot \mathbb{E}\left[\, W_i \mathbf{1}\left\{\, W_i \geq \tau_\theta \,\right\}\right] \leq \varepsilon' \cdot \mathbb{E}\left[\, W_i \,\right] \leq \varepsilon' \cdot \mathbb{E}\left[\, \max_{i \in U} W_i \,\right] = \varepsilon.$$

Thus, we obtain

$$\mathbb{E}\left[\, x(A_\theta, T) \,\right] > \mathbb{E}\left[\, x(A, T) \,\right] - \varepsilon,$$

a

Claim A.4.6. *Given any algorithm $A_\theta \in \mathcal{A}_\theta$, for any $\varepsilon > 0$, there exists sufficiently large m such that*

$$x(A_{\theta,m}, T) \geq x(A_\theta, T) - \varepsilon.$$

Proof. Consider any element i, and let $\tau_\theta(i)$ and $\tau_{\theta,m}(i)$ denote the threshold for this element under the algorithms A_θ and $A_{\theta,m}$ respectively.

$$x_i(A, T) = \Pr[\, i \text{ is feasible under } A \,]\mathbb{E}[\, W_i \mathbf{1}\{\, W_i \geq \tau(i) \,\}],$$

where feasibility means that element i is not blocked with respect to the set constraint $\mathcal{I}$ when the element i arrives. Note that this is independent from the event concerned with the value of W_i. Note that, as long as $W_j \notin (\tau_{\theta,m}(j), \tau_\theta(j))$ for all $j \neq i$, the two algorithms behave in the same way. Therefore, if i is feasible under A_θ, then it would also be feasible under $A_{\theta,m}$ conditioned on this event. It follows that

$$\Pr[\, i \text{ is feasible under } A_{\theta,m} \,] \geq \Pr[\, i \text{ is feasible under } A_\theta \,] - n\gamma/m.$$

Additionally, since $\tau_{\theta,m} \leq \tau_\theta$, we have

$$\mathbb{E}[\, W_i \mathbf{1}\{\, W_i \geq \tau_\theta \,\}] \leq \mathbb{E}[\, W_i \mathbf{1}\{\, W_i \geq \tau_{\theta,m} \,\}].$$

Thus, it follows that

$$\begin{aligned} x(A_{\theta,m}, T) &= \sum_{i \in U} x_i(A_{\theta,m}, T) \\ &\geq \sum_{i \in U} \left(x_i(A_\theta, T) - \mathbb{E}\left[W_i \mathbf{1}\left\{ W_i \geq \tau_\theta \right\} \right] \cdot n\gamma/m \right) \\ &= x(A_\theta, T) - \mathbb{E}\left[W_i \mathbf{1}\left\{ W_i \geq \tau_\theta \right\} \right] \cdot n\gamma/m. \end{aligned} \tag{A.15}$$

Now we take large enough m so that $m \geq \frac{n\gamma \cdot \mathbb{E}[\max_{i \in U} W_i]}{\varepsilon}$. Then we have

$$\frac{n\gamma}{m} \cdot \mathbb{E}\left[W_i \mathbf{1}\left\{ W_i \geq \tau_\theta \right\} \right] \leq \frac{n\gamma}{m} \mathbb{E}\left[W_i \right] \leq \frac{n\gamma}{m} \cdot \mathbb{E}\left[\max_{i \in U} W_i \right] \leq \varepsilon.$$

Hence from (A.15), we obtain

$$x(A_{\theta,m}, T) \geq x(A_\theta, T) - \varepsilon,$$

and it finishes the proof of the claim. a

Combining Claim A.4.5 and A.4.6, for any $\varepsilon > 0$, there exist sufficiently large (θ, m) such that

$$x(A_{\theta,m}, T) \geq x(A, T) - \varepsilon. \tag{A.16}$$

Note that such ε is well-defined once we fix the distributions $D_1, D_2, \ldots, D_n$ and the underlying set system $\mathcal{M}$.

Finally, we obtain

$$\begin{aligned}
\inf_{T\in\Delta(P_n)} \sup_{A\in\Delta(\mathcal{A})} \mathbb{E}\left[\, x(A,T)\,\right] &\geq \sup_{A\in\Delta(\mathcal{A})} \inf_{T\in\Delta(P_n)} \mathbb{E}\left[\, x(A,T)\,\right] && \text{(By Lemma A.3.1)} \\
&\geq \sup_{A\in\Delta(\mathcal{A}_{\theta,m})} \inf_{T\in\Delta(P_n)} \mathbb{E}\left[\, x(A,T)\,\right] && (\because \Delta(\mathcal{A}_{\theta,m}) \subseteq \Delta(\mathcal{A})) \\
&= \inf_{T\in\Delta(P_n)} \sup_{A\in\Delta(\mathcal{A}_{\theta,m})} \mathbb{E}\left[\, x(A,T)\,\right] && \text{(By Lemma A.3.2)} \\
&\geq \inf_{T\in\Delta(P_n)} \sup_{A\in\Delta(\mathcal{A})} \mathbb{E}\left[\, x(A,T)\,\right] - \varepsilon && \text{(By (A.16))}
\end{aligned}$$

In particular, in the third line we could apply the linear version of von Neumann's minimax theorem in Lemma A.3.2 since both P_n and $A_{\theta,m}$ are finite. This finishes the proof. a

A.4.9 Proof of Proposition 1.4.2

Proof. Let $R := \{(I,(x,y)) : \exists (x,y) \in \text{Supp}(X,Y), (I,(x,y)) = g(\textsc{Mec};\mathcal{M},x,y)\}$ be the range of the mechanism given the problem instance, *i.e.,* the set of tuples of the selected set I and the realized utility (x,y) of the random variables. Define $\textsc{Mec}'$ be the single-proposal mechanism with eligible set R. Assuming that the agent plays $g(\textsc{Mec};\mathcal{M},x,y)$ under $\textsc{Mec}'$, the outcome of $\textsc{Mec}$ and $\textsc{Mec}'$ are equivalent since $g(\textsc{Mec};\mathcal{M},x,y)$ belongs to the eligible set. Thus, it suffices to prove that $g(\textsc{Mec};\mathcal{M},x,y)$ is indeed the agent's best-response under $\textsc{Mec}'$, given the realization (x,y). Let $\sigma = g(\textsc{Mec};\mathcal{M},x,y)$ for simplicity. Suppose that there exists the agent's another strategy σ' such that if agent plays σ' under $\textsc{Mec}'$, he obtains strictly larger than what she obtains under the strategy σ. If $\sigma' \notin R$, then obviously σ' cannot be better than σ since σ' would be rejected by $\textsc{Mec}'$, thereby yielding zero utility for the agent. Suppose $\sigma' \in R$. This means that σ' is one of the best-response outcomes of $\textsc{Mec}$, *i.e.,* $\sigma' = g(\textsc{Mec};\mathcal{M},x',y')$ for some x' and y'. Therefore, the agent's utility under

σ' is

$$\sum_{i \in g(\text{MEC}; \mathcal{M}, x', y')} y_i.$$

Note, however, that

$$\sum_{i \in g(\text{MEC}; \mathcal{M}, x', y')} y_i \leq \sum_{i \in g(\text{MEC}; \mathcal{M}, x, y)} y_i,$$

since $g(\text{MEC}; \mathcal{M}, x, y)$ is the best-response given (x, y). Since RHS it the agent's utility under σ, this implies that σ' cannot yield strictly large utility for the agent, and it finishes the proof. a

Appendix B: Delegating to Multiple Agents

B.1 Numerical Experiments on the Tightness of Theorem 3.6.7

We here numerically verify that the lower bound on $\Pr[E]$ obtained in the proof of Theorem 3.6.7, *i.e.,* $\exp\left(-\frac{n^2}{2(n-1)^2}\right)$, is almost tight. Figure B.1 represents $\Pr[E]$ with respect to the value of n for some choices of k. To obtain the true value of probability of the event E, we simulate the principal and agents and compute the frequency of the event that each agent's utility is aligned, *i.e.,* proposing a solution that maximizes the principal's utility also maximizes the agent's utility, for every agent. To present some details in the experiment, we consider $n = \{5, 10, 20, 50, 100, 200, 300, 400\}$ and $k = \{10, 20, 100, 150, 1000\}$. For each pair of n and k, we run 30000 times to compute the probability that the event E happens. We repeat this procedure 60 times to obtain the variance for each setting, which is plotted as the shade in the figure. The main take-away here is that the lower bound and the average value of the quantity $\Pr[E]$ is indeed becoming tight as n increases. As a side note, one can observe that as n increases the precision error appears in each experiment (shade) becomes wider. This is because large n introduces more randomness in each experiment which requires more samples to stabilize its average, and thus we believe it not to be a meaningful phenomenon.

B.2 Further discussion on feasible mechanisms

Another reasonable choice of mechanism will be to require each agent submitting certain number, say two, of solutions, *i.e.,* multi-proposal mechanism. Under this mechanism, each agent might

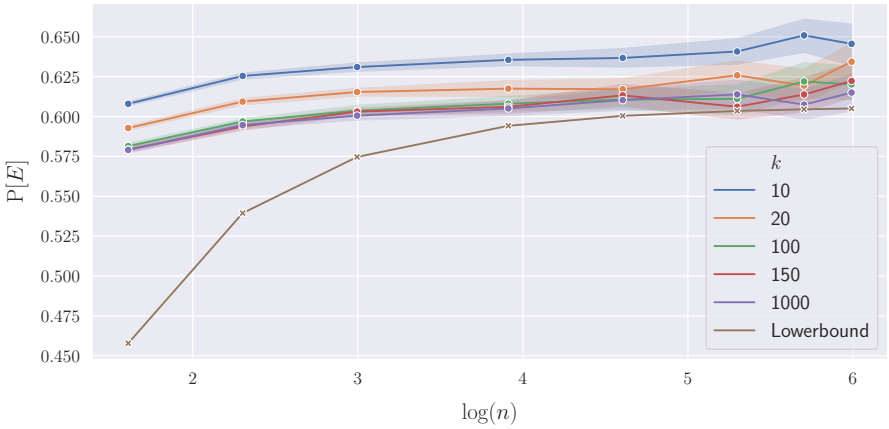


Figure B.1: Probability of the event E defined in (B.9). Lowerbound refers to the lower bound of the event E that we obtained in (B.15), and n and k denotes the number of agents and the number of solutions per agent. We note that x-axis is presented in the base of the natural number e.

select a pair of aggressive solution with large ex-post utility but smaller chance to be selected, and conservative solution with small ex-post utility but larger probability of winning, in order to balance these features of two solutions. More abstractly, one may consider a mechanism that requires each agent to signal some statistics of the observed solutions, *e.g.,* sample average or the number of eligible solutions, upon the solutions themselves. As astute readers may have noticed, in the complete information setting, however, the revelation principle style of result allows us to reduce any mechanism to single-proposal mechanism, so such generalization does not help. This is formalized in Theorem 3.4.11. Note that this reduction does not preclude any randomized mechanism.

B.3 Bayes Correlated equilibrium

In this section, we discuss that the mechanism presented in Theorem 3.6.7 also has an efficient Bayes correlated equilibrium. To be more specific, we are interested in the notion of Bayes correlated equilibrium [206]. Suppose that there exists a mediator who is able to access the realization of all the

solutions observed by the agents. Given the realization of all the solutions $\bar{\omega} = (\bar{\omega}_i)_{i\in[n]}$ observed by the agents, the mediator recommends a strategy $\psi_i : \Omega^* \mapsto \Sigma_i$ for each agent $i \in [n]$. It is important to note that the mediator recommends the strategy based on observing all the agents' solutions. Given the recommended strategy ψ_i and observed solution $\bar{\omega}_i$, agent i plays a strategy $\sigma_i(\bar{\omega}_i, \psi_i) : \Omega^* \times \Sigma_i \mapsto \Omega_i$. The procedure of recommending strategies is specified as a part of the mechanism M, and we denote by $f_{M,\sigma,\psi}(\bar{\omega})$ the interim allocation function, *i.e.,* the selected winner, of the mechanism M given $\psi(\bar{\omega})$ and $\sigma(\bar{\omega}, \psi)$.

Then, given $\psi_i(\bar{\omega})$ and $\bar{\omega}_i$, agent i's expected utility of playing σ_i is

$$u_i(\sigma_i(\psi_i, \bar{\omega}_i), \sigma_{-i}) = \mathbb{E}\left[\int_{\bar{\omega}_{-i}} y_i(f_{M,(\sigma_i(\psi_i(\bar{\omega}),\bar{\omega}_i),\sigma_{-i})}) f_{-i}(\bar{\omega}_{-i}) d\bar{\omega}_{-i}\right],$$

and agent i's total expected utility of playing σ_i is

$$u_i(\sigma_i, \sigma_{-i}) = \mathbb{E}_{\bar{\omega}_i}\left[u_i(\sigma_i(\psi_i, \bar{\omega}_i), \sigma_{-i})\right].$$

We now define a Bayes correlated equilibrium as follows.

Definition B.3.1 (Bayes correlated equilibrium)**.** We say that ψ_i induces a *Bayes correlated equilibrium* (BCE) if for any $i \in [n]$,

$$u_i(\psi_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i}),$$

for any other strategy σ_i', *i.e.,* there exists no incentive to deviate from the mediator's recommendation.

Compared to BNE, due to the existence of the additional signals provided by the mediator who

can observe all the solutions, both the principal and agents may enjoy the increased utility. In the following theorem, we prove that there exists a PIM such that it achieves asymptotically constant multiplicative approximation factor for some BCEs. We provide the proof in Appendix B.6.17

Theorem B.3.2. *Consider a symmetric and independent setting such that the distributions of the both principal's and agent's utility follow* $U[0,1]$*. Then, there exists an* $\left(\exp\left(\frac{n^2}{2(n-1)^2}\right), 0\right)$*-approximate PIM and corresponding BCE.*

B.4 Bulow-Klemperer-style result for prior-independent mechanism

In constructing efficient PIM, one standard way would be to construct PIM from BM by applying the Bulow-Klemperer-style result by [16], which considers a fictitious BM with $n+1$ agents to find a good upper bound on the efficiency of PIM with n agents. More precisely, they first show the existence of PIM (auction in their case) which always allocates the item to some bidders and then prove that such PIM is optimal among the mechanism that always allocates the item. Then, for any Bayesian mechanism M with an arbitrary reserve price, one can consider a fictitious auction with $n+1$ agents such that it runs M with the first n agents, and if it does not allocate the item, then it allocates the item to $n+1$-th agent. Given the observations above, one can conclude that the PIM with n agents defined above outperforms arbitrary BM with $n+1$ agents, and thus we can derive the approximation factor of PIM based on BM. This technique, however, does not work in our case as we cannot easily argue that MSPM that admits all the solutions induces the optimal utility for any mechanism that always adopts the winner. This is because we can consider a PIM that always admits all the solutions for some agents, but enforce some nontrivial (which does not always accept any solution) eligible sets for other agents. Indeed, this mechanism may outperform MSPM that admits all solutions, since it may motivate the

other agents with nontrivial eligible sets to propose more in favor of the principal, while maintaining the mechanism to always delegate to some agents.

B.5 Comparing multi-agent and single-agent delegation

To give a brief intuition in how one can reasonably compares the multi-agent and single-agent setting in terms of the principal, we start with an example.

Example B.5.1. Consider a task requester in the online labor market, *e.g.,* crowdsourcing platform. The task requester wants to solve a task by finding a single best solution for the task, and she is capable of recruiting five online workers. Due to the intrinsic volatility of the online platform, each worker is equipped with a similar level of skill, and is expected to exert a similar level of effort per some unit of payment. As such, each worker has the same *certainty equivalent* in coming up with a single solution, saying $\$10$. Because of the anonymity of the online labor market, each worker does not exactly obey the instruction of the task therein to answer the task, but instead propose a solution of his own interest, *i.e.,* misaligned payoffs. The task requester has a budget of $\$100$, and faces two options of (i) delegating to a single agent and (ii) delegating to five agents. The first option is to solely depend on a single worker, and the worker will sample 10 solutions and commit to one of them. By exploiting the second option, five workers will observe two solutions each and commit to one of them. Given the assumptions above, it is likely that the expected quality of a single solution will be similar, i.e. workers are symmetric. Which option should the principal take? As observed in our previous results, delegating to multiple workers is expected to be generally better, but to any extent?

In what follows, we answer this question in quantitative manner. More formally, we provide comparative statics given the reasonable comparison between the single-agent and multi-agent setting,

wherein our derivation of reasonable comparison follows from the intuition described in the example above. Note that we obtain the result of similar flavor in Theorem 3.5.3, but it only accommodates a reduction from a multi-agent instance to single-agent instance entirely to provide an approximation ratio of the multi-agent setting. In this section however, we cement this result by showing that given any single-agent instance with arbitrary mechanism, we can construct a reasonable multi-agent instance such that there exists a multi-agent mechanism with principal's utility at least that under the original single-agent instance. This result implies given any single-agent delegation problem instance, the principal can instead delegate to a reasonable set of multiple agents and obtain a superior utility. To this end, we first introduce how we construct a multi-agent instance from a single-agent instance.

Definition B.5.2 (Multi-agent correspondence)**.** Consider a single-agent problem with a distribution D, a number of solutions k, the agent's utility function y, and the principal's utility function x. Consider a multi-agent problem with $n \geq 2$ agents in which the principal has the same utility function x and each agent i has a distribution $D_{i,j} = D$ for $j \in [k_i]$, a number of solutions k_i such that $\sum_{i=1}^{n} k_i = k$, and the utility function $y_i = y$. We denote this multi-agent problem by multi-agent correspondence of the single-agent problem.

The above construction of multi-agent instance is reasonable. This is because to see whether delegating to the multi-agent indeed brings advantages, it is fair to compare under the assumption that the agents exert the same total amount of effort in the task in both the single-agent and the multi-agent problem. Since the number of solutions can be translated as the amount of effort exerted in the task, our definition of the multi-agent correspondence would play a reasonable role in comparing the single-agent and the multi-agent problem. We present the proof in Appendix B.6.18

Theorem B.5.3. *Given a single-agent problem P, denote its multi-agent correspondence P'. For any*

mechanism M under P, there exists an MSPM under P' such that the principal's utility of the MSPM under P' is at least that of M under P.

The main take-away is that the principal can actually increase (at least not decrease) her utility by recruiting more agents with similar utilities to delegate the task, although this does not quantify whether the multi-agent delegation largely improves upon the single-agent setting.

Remark B.5.4. As we discussed in Section 3.7, delegating to multiple agents induce a trade-off in examining the candidates in terms of the principal's resource, *e.g.,* time. Meanwhile, delegating to multiple agents already has its advantage in time consumption in the sense that it would take much smaller time for the entire process to be finished in the multi-agent delegation as each agent's task solving process is independent from each other and thus can be parallelized. We discuss more in the following example.

Example B.5.5. Back to Example B.5.1, it is indeed the case the task requester can simultaneously request five workers to solve the task and save her time. Assuming that obtaining a single solution takes around five minutes. Roughly speaking, single-agent delegation with 10 solutions requires 50 minutes, whereas delegating to five agents with two solutions would take 10 minutes. Afterward, the principal will examine the task. Suppose that the principal requires t minutes to examine a single solution. In the single-agent scenario, overall time consumption is $50+t$. Note that t is added back as the principal indeed has to examine the solution even though it is only one as the use of Bayesian mechanism (and thus decide whether or not to accept it) is essential in the single-agent setting. Meanwhile, the five agents scenario require $10 + 5t$. For any $t \leq 10 = 5 \cdot 2$, *i.e.,* when the principal can examine the solutions faster than twice of the time that each agent needs to find one solution, delegating to five agents still carries its own benefit even under the total examination. This will be even more true when

the production function of the worker is concave, *i.e.,* , when they need more time to find a solution as they've already found some solutions. In short, multi-agent delegation may have both time and efficiency gains to some extents.

B.6 Proofs

B.6.1 Proof of Theorem 3.4.11

Proof. Let R_i be the set of all the solutions belonging to agent i that can be selected by the mechanism M, over any feasible $\bar{\omega} \in \Omega^*$ and corresponding equilibrium strategies, and any random bits of the mechanism. Given a set of observed solutions by the agents $\bar{\omega} = \cup_{i\in[n]}\bar{\omega}_i \in \Omega^*$ and corresponding equilibrium strategy σ, consider any $\omega_M \in F_{M,\sigma}(\bar{\omega})$.

Suppose that ω_M belongs to agent i. Obviously, $\omega_M \in R_i$ holds due to our construction. Furthermore, in terms of agent $j \neq i$, there exists no solution $\omega' \in \bar{\omega}_j$ such that $x(\omega) > x(\omega_M)$ since if so, agent j will be selected in M by playing ω, *i.e.,* for any $\omega' \in \bar{\omega} \setminus \bar{\omega}_i$,

$$x(\omega_M) \geq x(\omega'). \tag{B.1}$$

Furthermore, if there exists another $\omega \in \bar{\omega}_i$ such that $x(\omega) \geq x(\omega')$ for all $\omega' \in \bar{\omega} \setminus \bar{\omega}_i$ and $y_i(\omega) > y_i(\omega_M)$, then agent i's utility becomes larger by playing ω, which contradicts the assumption that $\sigma_i = \omega_M$ is an equilibrium strategy. Hence, we conclude that whenever such ω exists,

$$y_i(\omega_M) \geq y_i(\omega). \tag{B.2}$$

Consider a MSPM with an arbitrary tie-breaking rule and eligible sets R_i for each agent i. Let this mechanism M'. Let σ' be arbitrary equilibrium strategies under M'. Let $\omega_{M'}$ be the winner of M'. We now consider two cases (i) when $\omega_{M'} \in \bar{\omega}_i$ and (ii) $\omega_{M'} \in \bar{\omega}_j$ for some $j \neq i$.

Firstly, assume that $\omega_{M'}$ belongs to agent i. This means that $\sigma'_i = \omega_{M'}$. Since σ'_i is an equilibrium strategy, there exists no solution $\omega' \in \bar{\omega} \setminus \bar{\omega}_i$ such that $x(\omega') > x(\sigma'_i)$ since if so, the agent who owns ω will be selected in M' by playing ω' which contradicts that σ'_i is an equilibrium strategy. Hence, we have $x(\sigma'_i) \geq x(\omega')$ for any $\omega' \in \bar{\omega} \setminus \bar{\omega}_i$. By (B.2), this implies that $y_i(\omega_M) \geq y_i(\sigma'_i) = y_i(\omega_{M'})$. By our assumption of Pareto optimal play, we further have

$$x(\sigma'_i) = x(\omega_{M'}) \geq x(\omega_M).$$

Otherwise, suppose that $\omega_{M'}$ belongs to agent $j \neq i$. In this case, suppose that $x(\omega_{M'}) < x(\omega_M)$. Then, agent i is strictly better off by playing ω_M. Hence, this is not equilibrium, and it completes the proof. a

B.6.2 Proof of Theorem 3.5.1

Proof. We mainly prove that the expected payoff of randomized single-proposal mechanism M is at least

$$\mathbb{E}\left[\, X_{(2)} 1[X_{(2)} \geq \tau] \,\right] + \mathbb{E}\left[\, \tau 1\left[X_{(1)} \geq \tau > X_{(2)}\right] \,\right].$$

We find a lower bound on the principal's utility in an arbitrary Nash equilibrium. Let $\bar{\omega} = \cup_{i\in[n]}\bar{\omega}_i$ be the entire set of solutions observed by the agents and $\sigma = (\sigma_1, \ldots, \sigma_n)$ be arbitrary Nash equilibrium. We separately consider two cases (i) for any $\omega \in \bar{\omega}$, $\omega \notin R$ and (ii) there exists $\omega \in \bar{\omega}$

such that $\omega \in R$.

Case (i). In this case, the utility of all agents is 0, and there does not exist a winner. Recall that for any Nash equilibrium, there exists no incentive for any agent to deviate from equilibrium strategies. We claim that $X_{(1)} < \tau$, where it concludes that the principal's utility is simply zero in this case. Suppose not, *i.e.,* $X_{(1)} \geq \tau$. Denote by ω^* one of the candidates which corresponds to the utility of $X_{(1)}$, and observed by agent t. If agent t proposes ω^*, his utility strictly increases from 0 to $y(\omega^*)$, which contradicts that the agents have played equilibrium strategies. Hence we conclude that $X_{(1)} < \tau$, and the principal's utility is 0.

Case (ii). We assume that agents play mixed strategies, *i.e.,* they decide on a probability distribution over the observed solutions. Formally, let $p_{i,j}$ be the probability that agent i proposes $\omega_{i,j}$. These probability distributions form a mixed Nash equilibrium. As agents play mixed strategies, for each $i \in [n]$ we have $\sum_{j=1}^{k_i} p_{i,j} = 1$. Define $S_i = \{\omega_{i,j} : p_{i,j} > 0\}$. For each agent i, let b_i be his worst candidate in terms of principal utility, *i.e.,* $b_i = \operatorname{argmin}_{\omega \in S_i} x(\omega)$. If there are multiple candidates with the worst utility, arbitrarily choose one. We now proceed with the following claim.

Claim B.6.1. *Denote the set of agents who have non-zero utility by A, then $|A| = 1$.*

Proof of the claim. Suppose not. Let t be an agent $t \in A$ such that for each agent $i \in A - \{t\}$, either

- $x(b_t) < x(b_i)$
- $x(b_t) = x(b_i)$, and principal prefers agent i over agent t using tie-breaking rule ρ,

i.e., t is the agent with the worst b_t who is least preferred according to the tie-breaking rule. Since agent $t \in A$ has non-zero utility and $b_t \in S_t$, proposing b_t also gives a non-zero utility to agent t. Let $\omega_{i,j} \in S_i$ be the proposed candidate by arbitrary agent $i \in A - \{t\}$. This is because if there exists a solution

with zero utility in S_t, then just removing b_t from S_t and distributing its assigned probability arbitrarily to other solutions will give a strictly higher utility for agent t. For any $i \in A \setminus \{t\}$ and $\omega_{i,j} \in S_i$, we observe that $x(\omega_{i,j}) \geq x(b_i)$. In addition, we have either $x(b_t) < x(\omega_{i,j})$, or $x(b_t) = x(b_i) = x(\omega_{i,j})$. Since the principal prefers i over t according to the tie-breaking rule ρ, the candidate proposed by agent t, *i.e.,* b_t, is dominated by the candidate proposed by agent i, *i.e.,* $\omega_{i,j}$. This implies that the principal never selects b_t as the winner, and it contradicts the fact that agent t has non-zero utility when proposing b_t. Therefore, we can conclude that $|A| = 1$. a

Suppose that agent $i^* \in A$ is the only agent with non-zero utility. In this case, all agents except agent i^* have a utility of zero. Due to the definition of Nash equilibrium, for any $\sigma_i = \omega$ for some $\omega \in \bar{\omega}_i$, $y_i(f_{M,(\sigma_i,\sigma_{-i})}) = 0$. This means that the principal never selects $\omega \in \bar{\omega}_i$ for agent $i \neq i^*$. Hence, for all $i \neq i^*, \omega_{i^*,j} \in S_{i^*}$, we have

$$x(\omega_{i^*,j}) \geq \max_{\omega \in \bar{\omega}_i} x(\omega) = X_{i,(1)}$$

Furthermore, for all $\omega_{i^*,j} \in S_{i^*}$ we have $x(\omega_{i^*,j}) \leq X_{i^*,(1)}$. Since for $n-1$ agents i ($i \neq i^*$) we have $X_{i,(1)} \leq x(\omega_{i^*,j})$, and $x(\omega_{i^*,j}) \leq X_{i^*,(1)}$ using the definition of $X_{(r)}$, for all $\omega_{i^*,j} \in S_{i^*}$ we get

$$X_{(n)} \leq X_{(n-1)} \leq ... \leq X_{(2)} \leq x(\omega_{i^*,j}) \leq X_{(1)}.$$

Also, in this case, all candidates in S_{i^*} should be eligible as they give non-zero utility to agent i^*. This further implies that for all $\omega_{i^*,j} \in S_{i^*}$, $\tau \leq x(\omega_{i^*,j}) \leq X_{(1)}$, and we have

$$\max\left(X_{(2)}, \tau\right) \leq x(\omega_{i^*,j}) \leq X_{(1)}.$$

This inequality shows that

- if $X_{(2)} \geq \tau$ then $x(\omega_{i^*,j}) \geq X_{(2)}$. Hence, in this case, the principal's expected utility is at least $X_{(2)}$ as $\mathbb{E}\left[\, f_{M,\sigma}(\bar{\omega})\,\right] = \sum_{j\in[k_{i^*}],\omega_{i^*,j}\in S_{i^*}} p_{i^*,j} x(\omega_{i^*,j}) \geq \sum_{j\in[k_{i^*}],\omega_{i^*,j}\in S_{i^*}} p_{i^*,j} X_{(2)} = X_{(2)}$.

- if $X_{(1)} \geq \tau \geq X_{(2)}$ then $x(\omega_{i^*,j}) \geq \tau$. Hence, in this case, the principal's expected utility is at least τ as $\mathbb{E}\left[\, f_{M,\sigma}(\bar{\omega})\,\right] = \sum_{j\in[k_{i^*}],\omega_{i^*,j}\in S_{i^*}} p_{i^*,j} x(\omega_{i^*,j}) \geq \sum_{j\in[k_{i^*}],\omega_{i^*,j}\in S_{i^*}} p_{i^*,j} \tau = \tau$.

Therefore, by combining the results of case (i) and case (ii) the expected utility of the principal is at least

$$\mathbb{E}\left[\, X_{(2)} 1\left[X_{(2)} \geq \tau\right]\,\right] + \mathbb{E}\left[\, \tau 1\left[X_{(1)} \geq \tau > X_{(2)}\right]\,\right]. \tag{B.3}$$

a

B.6.3 Proof of Theorem 3.5.3

Proof. Given the multi-agent delegation problem, we construct a single-agent instance and then use guarantees on the single-agent setting to prove our theorem. Consider a problem instance of multi-agent delegation where n agents propose solutions. In order to find solutions, agent i samples a solution from each of his k_i distributions $\{D_{i,j}\}_{j=1}^{k_i}$. Note that all solutions lie in Ω. We construct a single-agent problem instance where we denote the only agent by s. In the single-agent instance, for all $i \in [n]$, agent s samples k_i solutions, one from each of distributions $\{D_{i,j}\}_{j=1}^{k_i}$, hence $\sum_{i=1}^{n} k_i$ samples on aggregate.

Let ω be a solution that agent s observes from the distributions that belong to agent i. We represent this solution in single-agent instance with a pair (ω, i). Importantly, we define the solution

space Ω' in single-agent instance by $\Omega' := \Omega \times [n]$. In the constructed instance, the principal's utility of the solutions, *i.e.,* the function $x(.)$, does not change. Precisely, $x((\omega, i)) := x(\omega)$. However, we define $y((\omega, i))$, *i.e.,* agent s utility for ω when observed by distributions of agent i, to be $y((\omega, i)) := y_i(\omega)$.

According to [3], there exists an SPM with proper eligible set R under the constructed instance such that principal's expected utility is at least $\frac{1}{2}\mathbb{E}\left[X_{(1)}\right]$. We use eligible sets $R_i = R$ for all agents in the multi-agent instance. We claim that for any realization of solutions $\bar{\omega} = \cup_{i=1}^{n} \bar{\omega}_i$, the principal's utility in the multi-agent setting is at least that of the single-agent setting. This further implies that using eligible sets $R_i = R$, principal's expected utility in multi-agent settings is at least $\frac{1}{2}\mathbb{E}\left[X_{(1)}\right]$.

Consider a realization of solutions $\bar{\omega}$. Let (ω, i) be the winner in the single-agent setting. If $(\omega, i) = \perp$, then all observed solutions are not eligible. Formally, for all i in $[n]$, and j in $[k_i]$: $\omega_{i,j} \notin R$. In this case, there is no eligible solution in the multi-agent setting which indicates that the principal's utility in both settings is zero.

On the other hand, if $(\omega, i) \neq \perp$, let ω^* be the winner in the multi-agent instance, we claim that $x(\omega^*) \geq x(\omega)$. Note that (ω, i) is the winner of the single-agent instance so ω is the eligible solution that maximizes agent s utility by giving her the utility of $y\left((\omega, i)\right) = y_i(\omega)$.

Denote by σ_i the solution proposed by agent i in the multi-agent instance. As both $\sigma_i, \omega \in R$, and ω is the eligible solution that maximizes y_i, then $y_i(\omega) \geq y_i(\sigma_i)$. As a result, according to Assumption 3.4.9, $x(\sigma_i) \geq x(\omega)$. There are two cases

- ω^* is sampled by i, *i.e.,* $\omega^* = \sigma_i$. As shown above $x(\omega^*) = x(\sigma_i) \geq x(\omega)$.
- ω^* is sampled by another agent $j \neq i$. As the principal selects the best candidates among all agents, $x(\omega^*) \geq x(\sigma_i)$. Using the inequality $x(\sigma_i) \geq x(\omega)$ we further conclude that $x(\omega^*) \geq x(\omega)$.

This finishes the proof.

Finally, consider a correspondence between the multi-agent and single-agent setting such that the entire solutions in the multi-agent setting is sampled by the agent in the single-agent setting (as we did in the proof). Then, our proof in fact shows that the multi-agent problem can always bring higher (or equal) principal's utility than the single-agent problem equipped with any possible mechanism without money. This is because any mechanism in the single-agent setting can be reduced to the single-proposal mechanism by Lemma 1 in [3], and we prove that our construction of the mechanism above yields a higher (or equal) utility for the principal than the single-proposal mechanism in the single-agent setting.

a

B.6.4 Proof of Theorem 3.5.5

Proof. Consider a multi-agent instance where n agents propose solutions. If there exists an agent whose sampled solutions are always better than others agents' solutions, we call that agent a *super* agent. Formally, we say that agent i is super agent if for each $j \in [k_i]$, each $i' \in [n] \setminus \{i\}$, and each $j' \in [k_{i'}]$ when $\omega_{i,j} \sim D_{i,j}, \omega_{i',j'} \sim D_{i',j'}$ we have $x(\omega_{i,j}) \geq x(\omega_{i',j'})$.

When a multi-agent delegation problem instance has a super agent, then whatever the super agent proposes is taken as the winner as all other agents' proposed solutions are dominated by any solution of the super agent. As a result, the multi-agent instance reduces to a single-agent instance where only agent i exists as the single agent.

We now show that $(2, 0)$-approximation is tight for Bayesian mechanisms. We use the notion of a super agent and construct a multi-agent delegation problem where there exists a super agent so that

only his sampled solutions are important. We note that $X_{(1)}$ here corresponds to the solution observed by the super agent as other agents will never sample the best solution.

Assume this agent samples two solutions from distributions D_1, and D_2.

- D_1 always gives the same solution ω_1 to the agent such that $x(\omega_1) = 1, y(\omega_1) = 10$.
- D_2 gives the agent solution ω_{21} such that $x(\omega_{21}) = \frac{1}{\alpha}, y(\omega_{21}) = 5$ with probability of α while with probability of $1 - \alpha$, it gives solution ω_{22} to the agent such that $x(\omega_{22}) = \frac{\alpha}{1-\alpha}, y(\omega_{22}) = 5$. Note that α is a very small positive number.

There are two possibilities for sampled solutions by this agent either it is $\{\omega_1, \omega_{21}\}$, or it is $\{\omega_1, \omega_{22}\}$. We first consider $\mathbb{E}\left[\, X_{(1)} \,\right]$

$$\mathbb{E}\left[\, X_{(1)} \,\right] = \alpha\left(\frac{1}{\alpha}\right) + (1-\alpha)(1) = 2 - \alpha.$$

Now we consider the principal's policy. There are two cases

- principal accepts ω_1, *i.e.,* $\omega_1 \in R$ then as $y(\omega_1) > y(\omega_{21}) = y(\omega_{22})$, agent proposes ω_1 no matter what he samples from D_2. In this case, the principal's utility is 1.
- principal rejects ω_1, *i.e.,* $\omega_1 \notin R$. In this case, the principal accepts the other solution as it gives a non-zero utility to her. With probability of α, ω_{21} is sampled by the agent which gives the utility of $\frac{1}{\alpha}$ to the principal, however, ω_{22} is sampled with probability of $1 - \alpha$ which gives the utility of $\frac{\alpha}{1-\alpha}$ to the principal. Therefore, prinicpal's expected utility using this policy is $\alpha\left(\frac{1}{\alpha}\right) + (1-\alpha)\left(\frac{\alpha}{1-\alpha}\right) = 1 + \alpha$.

Denote by ALG the principal's utility using her policy. In both cases, the principal's expected utility

$\mathbb{E}[\,ALG\,] \leq 1+\alpha$. By putting $\alpha := \frac{\varepsilon}{3-\varepsilon}$, we have

$$(2-\varepsilon)\mathbb{E}[\,ALG\,] \leq (2-\varepsilon)(1+\alpha) = (2-\varepsilon)\left(1+\frac{\epsilon}{3-\epsilon}\right) = 2-\frac{\varepsilon}{3-\varepsilon} = 2-\alpha = \mathbb{E}\left[\,X_{(1)}\,\right].$$

Therefore, $(2-\varepsilon)\mathbb{E}[\,ALG\,] \leq \mathbb{E}\left[\,X_{(1)}\,\right]$. This indicates that the multiplicative approximation factor is at least $2-\varepsilon$, which finishes the proof. a

B.6.5 Proof of Theorem 3.5.8

Proof. Define $R_i = R = \{\omega : x(\omega) \geq r(n,k)\}$. Suppose that we run MSPM with the eligible set R. Due to Theorem 3.5.1, the principal's expected utility is at least

$$\begin{aligned}
\mathbb{E}\left[\,X_{(2)}1[X_{(2)} \geq \tau]\,\right] + \mathbb{E}\left[\,\tau 1[X_{(1)} \geq \tau > X_{(2)}]\,\right] &\geq \mathbb{E}\left[\,\tau 1[X_{(1)} \geq \tau]\,\right] \qquad \text{(B.4)}\\
&\geq \Pr\left[\,\exists i : X_{i,(1)} \geq r(n,k)\,\right] r(n,k)\\
&= \left(1-F(r(n,k))^{nk}\right) r(n,k)\\
&\geq \left(1-(r(n,k)^{\alpha})^{nk}\right) r(n,k)\\
&= r(n,k) - r(n,k)^{\alpha nk+1}, \qquad \text{(B.5)}
\end{aligned}$$

where the second inequality comes from our assumption on $F(x) \leq x^{\alpha}$. Plugging $r(n,k) = \left(\frac{1}{\alpha nk+1}\right)^{\frac{1}{\alpha nk}}$,[1] we can further obtain

$$\left(\frac{\alpha nk}{\alpha nk+1}\right)^{\frac{1}{\alpha nk}},$$

[1]Note that this choice of $r(n,k)$ comes from derivating (B.5) w.r.t. $r(n,k)$ by considering it as a function of αnk.

which yields the multiplicative approximation ratio of $\left(\frac{\alpha nk}{\alpha nk+1}\right)^{\frac{-1}{\alpha nk}}$. a

B.6.6 Proof of Theorem 3.5.10

Proof. In order to prove this theorem, we reuse the idea of super agent defined in the proof of Theorem 3.5.5. Before that, we claim that the principal cannot announce an eligible set other than Ω. The proof is based on straightforward arguments but we provide it to make it self-contained.

Claim B.6.2. *If $R_i \neq \Omega$ for some $i \in [n]$, then the additive price of anarchy can be arbitrarily close to L,* i.e., *for any $\varepsilon > 0$ there exists a problem instance such that* $\text{PoS}_a \geq L - \varepsilon$.

Proof of the claim. Since $R_i \neq \Omega$, there must be an element $\omega \in \Omega$ such that the principal does not admit. Assume that $x(\omega) = L - \varepsilon/2$, and agent i's distribution is a point mass on ω. Further, assume that other agents $j \neq i$ have distributions such that the principal's utility is at most $\varepsilon/2$ for any point in their support. Then, our MSPM with the given eligible sets only accepts some solutions from $j \neq i$, which results in the principal's utility of at most $\varepsilon/2$. Meanwhile, the optimal utility will be to select agent i's solution ω under which the principal can obtain $L - \varepsilon$. Hence, the additive approximation factor is $L - \varepsilon/2 - \varepsilon/2 = L - \varepsilon$, and it completes the proof. a

Due to the claim above, the principal should announce no eligible set in order to derive a nontrivial approximation ratio. Consider a multi-agent instance such that there exists a super agent. Because the principal does not have any information regarding the distribution of agents, she cannot announce any eligible sets which implies that she accepts all solutions and takes the best one among the proposed solutions. As a result, whatever the super agent proposes is selected by the principal. Assume the super agent samples two solutions from distributions D_1 and D_2. D_1 always gives the same solution ω_1 to the agent such that $x(\omega_1) = L - \varepsilon/2, y(\omega_1) = 5$. D_2 always gives the agent solution ω_2 such that

$x(\omega_2) = \varepsilon/2, y(\omega_2) = 10$. We obtain

$$\mathbb{E}\left[\, X_{(1)} \,\right] = \mathbb{E}\left[\, \max\{\varepsilon/2, L - \varepsilon/2\} \,\right] = L - \varepsilon/2.$$

The super agent proposes ω_2 as it will not be rejected by the principal and gives a higher utility to him, which means that the principal's utility is $x(\omega_2) = \varepsilon/2$. Therefore, the additive approximation factor of this mechanism is $\mathbb{E}\left[\, X_{(1)} \,\right] - x(\omega_2) = L - \varepsilon$. This finishes the proof. a

B.6.7 Proof of Theorem 3.5.11

Proof. For ease of presentation, assume that $L = 1$. By our assumption on symmetry and since each agent independently samples solutions, $X_{i,(1)}$ for $i \in [n]$ are identical and independent. Then, the result directly follows from the following theorem.

Lemma B.6.3 ([207]). *Let $X_1, \ldots, X_n$ be i.i.d. random variables supported on $[0, 1]$. Given an integer s such that $1 \leq s \leq n - 1$, let Y_s and Y_{s+1} be s-th and $s + 1$-th order statistics of $X_1, \ldots, X_n$. Then, the following holds.*

$$\mathbb{E}\left[\, Y_{s+1} - Y_s \,\right] \leq \binom{n}{s} \left(\frac{s}{n}\right)^s \left(1 - \frac{s}{n}\right)^{n-s}.$$

Plugging $s = n - 1$ yields the desired result for $L = 1$. By scaling the random variables, we can easily obtain the general of upper bound $L(1 - 1/n)^{n-1}$.

To prove that this upper bound is tight upper bound for $\mathbb{E}\left[\, X_{(1)} - X_{(2)} \,\right]$, consider a symmetric problem instance equipped with a Bernoulli distribution with parameter p, *i.e.,* $X_{i,j} \sim \text{Bernoulli}(p)$. In this case, $X_{i,(1)}$ is also a Bernoulli distribution with parameter $q = 1 - (1 - p)^k$. Furthermore, we

have

$$\mathbb{E}\left[\, X_{(1)} - X_{(2)} \,\right] = \binom{n}{n-1}(1-q)^{n-1}q.$$

If we set $q = 1-(1-p)^k = \frac{1}{n}$, which is equivalent to $p = 1-(1-1/n)^{1/k}$, then we have the tight bound of $(1-1/n)^{n-1}$. By scaling this problem instance, we finish the proof.

a

B.6.8 Proof of Lemma 3.5.12

Proof. We separately prove two statements using similar techniques. Let F and f be the c.d.f. and p.d.f. of D, respectively.

Part (1). Rearranging $\Delta_{k,n} \leq \Delta_{k,n-1}$, we need to show that

$$\mathbb{E}\left[\, X_{n-k+1:n} - X_{n-k:n-1} \,\right] \leq \mathbb{E}\left[\, X_{n-k:n} - X_{n-k-1:n-1} \,\right].$$

Due to [208], LHS and RHS can be represented as follows.

$$\mathbb{E}\left[\, X_{n-k+1:n} - X_{n-k:n-1} \,\right] = \binom{n-1}{n-k}\int_{-\infty}^{\infty} F^{n-k}(x)(1-F(x))^k dx.$$
$$\mathbb{E}\left[\, X_{n-k:n} - X_{n-k-1:n-1} \,\right] = \binom{n-1}{n-k-1}\int_{-\infty}^{\infty} F^{n-k-1}(x)(1-F(x))^{k+1} dx.$$

Plugging into the inequality and rearranging it, we need to show that

$$\binom{n-1}{n-k-1}\left(\int_{-\infty}^{\infty} F^{n-k-1}(x)(1-F(x))^k\left(\frac{n}{n-k}F(x)-1\right)dx\right)\leq 0. \tag{B.6}$$

Define $G(t)=\int_{-\infty}^{t} F^{n-k-1}(x)(1-F(x))^{k-1}\left(\frac{n-1}{n-k}F(x)-1\right)f(x)dx$. Obviously, $\frac{n-1}{n-k}F(x)-1$ inside the integrand becomes positive for $x>x_0$ and negative otherwise for some $x_0\in\mathbb{R}$. This implies that $G(t)\leq G(\infty)$ for any $t\in\mathbb{R}$. Let $u=F(x)$ and by change of variables, we have

$$\begin{aligned}
G(\infty) &= \int_{-\infty}^{\infty} u^{n-k-1}(1-u)^{k-1}\left(\frac{n}{n-k}u-1\right)du \\
&= \frac{n}{n-k}B(n-k+1,k)-B(n-k,k) \\
&= \frac{n}{n-k}\frac{(n-k)!(k-1)!}{n!}-\frac{(n-k-1)!(k-1)!}{(n-1)!}=0
\end{aligned}$$

where $B(\cdot,\cdot)$ denotes the beta function. Hence, we have $G(t)\leq 0$ for any $t\in\mathbb{R}$. Now we use the following lemma.

Lemma B.6.4 ([209]). *Given two functions $g:\mathbb{R}\mapsto[a,b]$ and $h:\mathbb{R}\mapsto[a,b]$ and h is nonnegative, then the following holds.*

1. *If $\int_{-\infty}^{t} g(x)dx\leq 0$ for any $t\in\mathbb{R}$, then for any nonincreasing function $h:\mathbb{R}\mapsto\mathbb{R}$, we have $\int_{-\infty}^{t} g(x)h(x)dx\leq 0$ for any $t\in\mathbb{R}$.*

2. *If $\int_{-\infty}^{t} g(x)dx\geq 0$ for any $t\in\mathbb{R}$, then for any nondecreasing function $h:\mathbb{R}\mapsto\mathbb{R}$, we have $\int_{-\infty}^{t} g(x)h(x)dx\geq 0$ for any $t\in\mathbb{R}$.*

Plugging $g(x)$ to be the integrand of $G(x)$ and $h(x)=(\frac{f(x)}{1-F(x)})^{-1}$ to the first part of the above lemma, we conclude that (B.6) holds since $h(t)$ is nonincreasing due to the assumption on MHR. This

completes the first part of the lemma.

Part (2). Now consider the MRHR distribution case. Rearranging $\delta_{k,n} \leq \delta_{k,n-1}$, we need to show that

$$\mathbb{E}\,[\,X_{k+1:n} - X_{k+1:n-1}\,] \leq \mathbb{E}\,[\,X_{k:n} - X_{k:n-1}\,].$$

Again due to [208], we have

$$\begin{aligned}
\mathbb{E}\,[\,X_{k+1:n} - X_{k+1:n-1}\,] &= -\binom{n-1}{k} \int_{-\infty}^{\infty} F^{k+1}(x)(1-F(x))^{n-k-1} dx. \\
\mathbb{E}\,[\,X_{k:n} - X_{k:n-1}\,] &= -\binom{n-1}{k-1} \int_{-\infty}^{\infty} F^{k}(x)(1-F(x))^{n-k} dx.
\end{aligned}$$

Plugging into the inequality and rearranging it, it suffices to show that

$$\frac{(n-1)!}{(k-1)!(n-k)!} \int_{-\infty}^{\infty} F^{k}(x)(1-F(x))^{n-k-1} \Big(\frac{n}{k} F(x) - 1\Big) dx \geq 0.$$

Define $G'(t) = \int_t^{\infty} F^{k-1}(x)(1-F(x))^{n-k-1}\big(\frac{n}{k}F(x) - 1\big) f(x) dx$. Observe that $\big(\frac{n}{k}\big)F(x) - 1$ becomes positive for $x > x_0$ and negative otherwise for some $x_0 \in \mathbb{R}$. Hence, $G'(t) \geq G'(-\infty)$. Now we show that $G'(-\infty) \geq 0$. Using change of variable with $F(x) = u$, we have

$$\begin{aligned}
G'(-\infty) &= \int_{-\infty}^{\infty} u^{k-1}(1-u)^{n-k-1} \Big(\frac{n}{k} u - 1\Big) du \\
&= \frac{n}{k} B(k+1, n-k) - B(k, n-k) \\
&= \frac{n}{k} \frac{k!(n-k-1)!}{n!} - \frac{(k-1)!(n-k-1)!}{(n-1)!} = 0,
\end{aligned}$$

and it concludes that $G'(t) \geq 0$ for any $t \in \mathbb{R}$. Now, plugging $g(x)$ to be the integrand of $G'(x)$ and $h(x) = (\frac{f(x)}{F(x)})^{-1}$ to the second part of Lemma B.6.4, we conclude that (B.6) holds since $h(t)$ is nondecreasing due to the assumption on MRHR. This completes the first part of the lemma.

a

B.6.9 Proof of Lemma 3.5.14

Proof. Let $f(x)$ and $F(x)$ be p.d.f. and c.d.f. of the distribution $X_i \sim D$, and let $g(x)$ and $G(x)$ be that of r-th order statistics, respectively. Then, we have

$$1 - G(x) = \sum_{j=0}^{r-1} \binom{n}{j} F(x)^j (1 - F(x))^{n-j}$$

$$g(x) = \frac{n!}{(r-1)!(n-r)!} f(x) F(x)^{r-1} (1 - F(x))^{n-r}.$$

Hence, its hazard rate can be obtained as

$$\begin{aligned} h(x) &= \frac{\frac{n!}{(r-1)!(n-r)!} f(x) F(x)^{r-1} (1 - F(x))^{n-r}}{\sum_{j=0}^{r-1} \binom{n}{j} F(x)^j (1 - F(x))^{n-j}} \\ &= \frac{n!}{(r-1)!(n-r)!} \frac{f(x)}{1 - F(x)} \frac{F(x)^{r-1} (1 - F(x))^{n-r+1}}{\sum_{j=0}^{r-1} \binom{n}{j} F(x)^j (1 - F(x))^{n-j}} \\ &= \frac{n!}{(r-1)!(n-r)!} \frac{f(x)}{1 - F(x)} \frac{1}{\sum_{j=0}^{r-1} \binom{n}{j} \left(\frac{1-F(x)}{F(x)}\right)^{r-j-1}}. \end{aligned}$$

Since $f(x)/(1 - F(x))$ is nondecreasing by MHR property of D and $1 - F(x)/F(x)$ is nonincreasing function of x, we conclude that $h(x)$ is nondecreasing function of x. a

B.6.10 Proof of Theorem 3.5.15

Proof. Recall that the additive approximation factor is upper bounded by $\mathbb{E}\left[X_{(1)} - X_{(2)} \right]$ by Corollary 3.5.9. As we discussed, the proof mainly follows from Lemma 3.5.12 on the expected difference between two consecutive order statistics. First, by Lemma 3.5.14, $X_{i,(1)}$ has MHR distribution for $i \in [n]$. Now, denote by $\mathbb{E}\left[X_{(1):n} - X_{(2):n} \right]$ the additive approximation factor in the problem setting when there exists n agents. By repeating Lemma 3.5.12 with $k = 1$, we can obtain the following inequalities.

$$\mathbb{E}\left[X_{(1):n} - X_{(2):n} \right] \leq \mathbb{E}\left[X_{(1):2} - X_{(2):2} \right] = \mathbb{E}\left[|A - B| \right],$$

where A and B are two i.i.d. random variables of which the endowed distribution is equivalent to that of $X_{i,(1)}$. We now use the following inequality.[2]

Lemma B.6.5 (Theorem 5 in [210])**.** *Let A, B be i.i.d. random variables with mean μ and variance σ^2. For any p, q such that $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, suppose that $\mathbb{E}\left[|A - \mu|^p \right] < \infty$. Then, the following inequality holds.*

$$\mathbb{E}\left[|A - B| \right] \leq \frac{2}{(q+1)^{1/q}} \mathbb{E}\left[|A - \mu|^p \right]^{1/p}.$$

Plugging $p = q = 2$ into the lemma, we obtain

$$\mathbb{E}\left[|A - B| \right] \leq \sqrt{\frac{4\mathrm{Var}(A)}{3}}.$$

[2] Weaker bound of $\mathbb{E}\left[|A - B| \right] \leq \sqrt{\mathrm{Var}(a)}$ can be obtained by simply rewriting it as $\mathbb{E}\left[\sqrt{(A-B)^2} \right]$ and applying Jensen's inequality.

In case of uniform distribution $U[0,1]$, it is widely known that $X_{i,(1)} \sim Beta(k,1)$. Hence, we conclude that the additive approximation factor is bounded above by

$$\sqrt{\frac{4k}{3(k+1)^2(k+2)}} \approx \frac{1.155}{k}.$$

a

B.6.11 Proof of Lemma 3.5.16

Proof. We separately prove two cases.

Part (1). Rearrange $(n+1)\Delta_{k,n} = \Delta_{k,n} + n\Delta_{k,n} \leq n\Delta_{k,n-1}$ and divide it by $n-1$,, we need to show that

$$\frac{\Delta_{k,n}}{n} + \mathbb{E}\left[X_{n-k+1:n} - X_{n-k:n-1} \right] \leq \mathbb{E}\left[X_{n-k:n} - X_{n-k-1:n-1} \right].$$

Again to [208], each term can be represented as follows.

$$\begin{aligned}
\frac{\Delta_{k,n}}{n} &= \frac{1}{n}\binom{n}{k}\int_{-\infty}^{\infty} F^{n-k}(x)(1-F(x))^k dx \\
\mathbb{E}\left[X_{n-k+1:n} - X_{n-k:n-1} \right] &= \binom{n-1}{n-k}\int_{-\infty}^{\infty} F^{n-k}(x)(1-F(x))^k dx. \\
\mathbb{E}\left[X_{n-k:n} - X_{n-k-1:n-1} \right] &= \binom{n-1}{n-k-1}\int_{-\infty}^{\infty} F^{n-k-1}(x)(1-F(x))^{k+1} dx.
\end{aligned}$$

Plugging into the inequality and rearranging it, we need to show that

$$\binom{n-1}{n-k-1}\left(\frac{1}{n-k}\int_{-\infty}^{\infty}F^{n-k}(x)(1-F(x))^k dx+\int_{-\infty}^{\infty}F^{n-k-1}(x)(1-F(x))^k\left(\frac{n}{n-k}F(x)-1\right)dx\right)$$
$$\leq 0.$$

We can further expand it to be

$$\binom{n-1}{n-k-1}\left(\int_{-\infty}^{\infty}F^{n-k-1}(x)(1-F(x))^k\left(\frac{n+1}{n-k}F(x)-1\right)dx\right)\leq 0. \tag{B.7}$$

Define

$$G(t)=\int_{-\infty}^{t}F^{n-k-1}(x)(1-F(x))^k\left(\frac{n+1}{n-k}F(x)-1\right)f(x)dx$$

Obviously, $\frac{n+1}{n-k}F(x)-1$ inside the integrand becomes positive for $x > x_0$ and negative otherwise for some $x_0 \in \mathbb{R}$. This implies that $G(t) \leq G(\infty)$ for any $t \in \mathbb{R}$. Let $u = F(x)$ and by change of variables, we have

$$\begin{aligned}
G(\infty) &= \int_{-\infty}^{\infty}u^{n-k-1}(1-u)^k\left(\frac{n+1}{n-k}u-1\right)du \\
&= \frac{n+1}{n-k}B(n-k+1,k+1)-B(n-k,k+1) \\
&= \frac{n+1}{n-k}\frac{(n-k)!k!}{(n+1)!}-\frac{(n-k-1)!k!}{n!}=0
\end{aligned}$$

where $B(\cdot,\cdot)$ denotes the beta function. Hence, we have $G(t) \leq 0$ for any $t \in \mathbb{R}$. Plugging $g(x)$ to be the integrand of $G(x)$ and $h(x) = 1/f(x)$ to the first part of Lemma B.6.4, we conclude that (B.7)

holds since $h(t)$ is nonincreasing due to the assumption on $f(x)$. This completes the first part of the lemma.

Part (2). Rearranging $(n+1)\delta_{k,n} \leq n\delta_{k,n-1}$, we need to show that

$$\frac{1}{n}\mathbb{E}\left[\, X_{k+1:n} - X_{k:n} \,\right] + \mathbb{E}\left[\, X_{k+1:n} - X_{k+1:n-1} \,\right] \leq \mathbb{E}\left[\, X_{k:n} - X_{k:n-1} \,\right].$$

Due to [208], we obtain the following equations.

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\left[\, X_{k+1:n} - X_{k:n} \,\right] &= \frac{1}{n}\binom{n}{n-k}\int_{-\infty}^{\infty} F^k(x)(1-F(x))^{n-k}dx \\
\mathbb{E}\left[\, X_{k+1:n} - X_{k+1:n-1} \,\right] &= -\binom{n-1}{k}\int_{-\infty}^{\infty} F^{k+1}(x)(1-F(x))^{n-k-1}dx \\
\mathbb{E}\left[\, X_{k:n} - X_{k:n-1} \,\right] &= -\binom{n-1}{k-1}\int_{-\infty}^{\infty} F^{k}(x)(1-F(x))^{n-k}dx
\end{aligned}$$

Plugging into the inequality and rearranging it, we need to show that

$$\frac{(n-1)!}{(k-1)!(n-k)!}\left(-\frac{1}{k}\int_{-\infty}^{\infty} F^k(x)(1-F(x))^{n-k}dx + \int_{-\infty}^{\infty} F^k(x)(1-F(x))^{n-k-1}\left(\frac{n}{k}F(x)-1\right)\right) \geq 0,$$

which is equivalent to

$$\frac{(n-1)!}{(k-1)!(n-k)!}\left(\int_{-\infty}^{\infty} F^k(x)(1-F(x))^{n-k-1}\left(\frac{n+1}{k}F(x)-\frac{k+1}{k}\right)\right) \geq 0.$$

Define $G'(t) = \int_t^{\infty} F^{k-1}(x)(1-F(x))^{n-k-1}\left(\frac{n+1}{k}F(x)-\frac{k+1}{k}\right)f(x)dx$. Observe that $\frac{n+1}{k}F(x)-\frac{k+1}{k}$ becomes positive for $x > x_0$ and nonpositive otherwise for some $x_0 \in \mathbb{R}$. Hence, $G'(t) \geq G'(-\infty)$.

Now we show that $G'(-\infty) \geq 0$. Using change of variable with $F(x) = u$, we have

$$\begin{aligned} G'(-\infty) &= \int_{-\infty}^{\infty} u^{k-1}(1-u)^{n-k-1}\left(\frac{n+1}{k}u - \frac{k+1}{k}\right)du \\ &= \frac{n+1}{k}B(k+1, n-k) - \frac{k+1}{k}B(k, n-k) \\ &= \frac{n+1}{k}\frac{k!(n-k-1)!}{n!} - \frac{k+1}{k}\frac{(k-1)!(n-k-1)!}{(n-1)!} \\ &= \frac{(k-1)!(n-k-1)!}{(n-1)!}\left(\frac{n+1}{n} - \frac{k+1}{k}\right) \leq 0, \end{aligned}$$

where the last inequality follows from $k \leq n-1$. Hence we conclude that $G'(t) \geq 0$ for any $t \in \mathbb{R}$. Now, plugging in $g(x)$ to be the integrand of $G'(x)$ and $h(x) = f(x)^{-1}$ to the second part of Lemma B.6.4, we finish the proof. a

B.6.12 Proof of Lemma 3.5.17

Proof. We separately prove the following two cases.

Part (1). Let $f(x)$ and $F(x)$ be p.d.f. and c.d.f. of the distribution $X_i \sim D$, and let $g(x)$ be p.d.f. of n-th order statistics, respectively. Then, we have

$$g(x) = \frac{n!}{(n-1)!(n-n)!}f(x)F^{n-1}(x)(1-F(x))^{n-n} = nf(x)F(x)^{n-1}.$$

Since both $f(x)$ and $F(x)$ are nondecreasing on x, we finish the proof.

Part (2). Let $g(x)$ be p.d.f. of first order statistics. Then, we have

$$g(x) = \frac{n!}{(1-1)!(n-1)!}f(x)F^{1-1}(x)(1-F(x))^{n-1} = nf(x)(1-F(x))^{n-1}.$$

Since both $1 - F(x)$ and $f(x)$ are nonincreasing on x, we finish the proof. a

B.6.13 Proof of Theorem 3.5.18

Proof. Recall that the additive approximation factor is upper bounded by $\mathbb{E}\left[X_{(1)} - X_{(2)} \right]$ by Corollary 3.5.9. As we discussed above, the proof mainly follows from Lemma 3.5.16 on the expected difference between two consecutive order statistics. By Lemma 3.5.17, we know that $X_{i,(1)}$ has p.d.f. which is nondecreasing on x. Now, denote by $\mathbb{E}\left[X_{(1):n} - X_{(2):n} \right]$ the additive approximation factor in the problem setting when there exists n agents. By repeating Lemma 3.5.16 with $k = 1$, we have the following series of inequalities.

$$
\begin{aligned}
(n+1)\mathbb{E}\left[X_{(1):n} - X_{(2):n} \right] &\leq n\mathbb{E}\left[X_{(1):n-1} - X_{(2):n-1} \right] \\
n\mathbb{E}\left[X_{(1):n-1} - X_{(2):n-1} \right] &\leq (n-1)\mathbb{E}\left[X_{(1):n-2} - X_{(2):n-2} \right] \\
&\vdots \\
4\mathbb{E}\left[X_{(1):3} - X_{(2):3} \right] &\leq 3\mathbb{E}\left[X_{(1):2} - X_{(2):2} \right].
\end{aligned}
$$

By multiplying the inequalities, we obtain

$$
\mathbb{E}\left[X_{(1):n} - X_{(2):n} \right] \leq \frac{3}{n+1}\mathbb{E}\left[|A - B| \right],
$$

where A and B are two i.i.d. random variables of which the endowed distribution is equivalent to that of $X_{i,(1)}$. Similarly from the proof of Theorem 3.5.15, we can further obtain

$$
\mathbb{E}\left[|A - B| \right] \leq \sqrt{\frac{4}{3}\mathrm{Var}(A)},
$$

and it completes the proof. In case of uniform distribution $U[0,1]$, it is widely known that $X_{i,(1)} \sim Beta(k,1)$. Hence, we conclude that the additive approximation factor is upper bounded by

$$\frac{3}{n+1}\sqrt{\frac{4k}{3(k+1)^2(k+2)}} \approx \frac{3.464}{k(n+1)}.$$

a

B.6.14 Proof of Theorem 3.6.4

Proof. The proof consists of two parts. First, we prove that there exists a problem instance such that proposing a solution that minimizes the principal's utility constructs an equilibrium, which leads to the principal's utility of $\mathbb{E}\left[\max_{i\in[n]} X_{i,(k)}\right]$. Next, we analyze the bound on this quantity under the uniform distribution.

Worst-case instance. Consider n symmetric agents that find solutions ω such that $x(\omega) = x_0$ is sampled from a continuous distribution with c.d.f. of $F(x_0)$. Each agent samples k solutions. For example, agent i samples $\omega_{i,1}, \omega_{i,2}, ..., \omega_{i,k}$. For an observed solution $\omega_{i,j}$, define agents' utility function y as a function of $x(\omega_{i,j})$ as follows

$$y(\omega_{i,j}) = \left(1 - \left((1 - F(x(\omega_{i,j})))^k\right)^{n-1}\right)^{-2}.$$

Under this setting, we claim that there exists a BNE such that each agent, among all of his sampled solutions, proposes the solution which minimizes the principal's utility. In order to show this claim, we prove that assuming other agents propose solutions with the lowest principal's utility, best-response for agent i is also to submit the solution which minimizes the principal's utility. Denote others' strategy

by σ_{-i}.

$$\mathbb{E}\left[\, u_i(\omega, \sigma_{-i}) | x(\omega) = x_0 \,\right] = \Pr\left[\, \omega \text{ is winner} | \sigma_{-i}, x(\omega) = x_0 \,\right] y(\omega)$$

As all other agents propose solutions with the lowest principal's utility, the probability that ω is chosen as the winner is the probability that all agents observe at least one solution ω' such that $x(\omega') \leq x(\omega) = x_0$. Note that tie does not happen as F is a continuous function. Formally,

$$\Pr\left[\, \omega \text{ is winner} | \sigma_{-i}, x(\omega) = x_0 \,\right] = \Big(1 - (1 - F(x_0))^k\Big)^{n-1}.$$

Plugging in the above equation, we have

$$\mathbb{E}\left[\, u_i(\omega, \sigma_{-i}) | x(\omega) = x_0 \,\right] = \Big(1 - (1 - F(x_0))^k\Big)^{n-1} y(\omega) = \frac{1}{\Big(1 - (1 - F(x_0))^k\Big)^{n-1}}$$

This indicates that agent utility is a decreasing function of x_0 which further implies that agent i best response is to propose solution ω such that $\omega = \operatorname{argmin}_{\omega \in \bar{\omega}_i} x(\omega)$. We prove that agent i best-response is to propose the solution with the lowest principal's utility, which concludes that this set of strategies forms a BNE. Hence in this case, the expected utility of the principal is $\mathbb{E}\left[\, \max_{i \in [n]} X_{i,(k)} \,\right]$.

Computing lower bound of $\textsc{PoA}_a$ for uniform distribution. Now we derive an upper bound for this quantity on the uniform distribution $U[0,1]$ setting, *i.e.,* $X_{i,j} \sim U[0,1]$. Note that $X_{i,(k)}$ follows the beta distribution with parameter $(1, k)$. By Jensen's inequality,

$$e^{t\left(\mathbb{E}\left[\max_{i \in [n]} X_{i,(k)}\right] - 1/(k+1)\right)} \leq \mathbb{E}\left[\, e^{t \cdot \left(\max_{i \in [n]} X_{i,(k)} - 1/(k+1)\right)} \,\right] \leq \sum_{i=1}^{n} \mathbb{E}\left[\, e^{t(X_{i,(k)} - 1/(k+1))} \,\right]. \tag{B.8}$$

Now we use the fact that beta distribution is subgaussian with certain choices of parameters.

Theorem B.6.6 (Theorem 4 in [211]). *The beta distribution with parameter α, β is $\frac{1}{4(\alpha+\beta)+2}$-subgaussian.*

Plugging $\alpha = 1$ and $\beta = k$, we obtain that $X_{i,(k)}$ is $\frac{1}{4k+6}$-subgaussian. Due to the definition of subgaussianity, we have

$$\mathbb{E}\left[\, e^{t(X_{i,(k)})} \,\right] \leq e^{t^2/(8k+12)}.$$

Plugging back to (B.8) and rearranging, we obtain

$$\mathbb{E}\left[\max_{i\in[n]} X_{i,(k)}\right] \leq \frac{1}{k+1} + \inf_{t>0} \frac{1}{t} \log\left(n \cdot e^{t^2/(8k+12)}\right) = \frac{1}{k+1} + \sqrt{\frac{\log n}{2k+3}}.$$

Since $\mathbb{E}\left[\, X_{\max} \,\right] = \frac{nk}{nk+1}$, we conclude that the additive approximation factor is at least

$$\textsc{PoA}_a \geq \frac{nk}{nk+1} - \frac{1}{k+1} - \sqrt{\frac{\log n}{2k+3}} \geq \frac{k-1}{k+1} - \sqrt{\frac{\log n}{2k+3}}.$$

This concludes that there exists no mechanism such that $\textsc{PoA}_a < \frac{k-1}{k+1} - \sqrt{\frac{\log n}{2k+3}}$. We note that by fixing n and increasing k sufficiently, we can make this bound arbitrarily close to 1, and we finish the proof. a

B.6.15 Proof of Theorem 3.6.7

Proof. The proof consists of two main parts. First, we construct an event E such that there exists an $\Pr\left[\, E^c \,\right]$-approximate BNE which yields the optimal principal's utility. Then, we analyze the probability of the event E happening.

Existence of approximate BNE. Consider an MSPM with no eligible sets and a deterministic tie-breaking rule ρ. Define an event $E_i(\cdot)$ as follows.

$$E_i(\{x(\omega), y(\omega)\}_{\omega\in\bar{\omega}_i}) = \left\{\bar{\omega}_i : \operatorname*{argmax}_{\omega\in\bar{\omega}_i} x(\omega)^{k(n-1)}y(\omega) = \operatorname*{argmax}_{\omega\in\bar{\omega}_i} x(\omega)\right\}.$$

For notational simplicity, let $x(A) = \{x(\omega)\}_{\omega\in A}$ and similarly for $y(A)$. We often overwrite

$$E_i(\{x(\omega), y(\omega)\}_{\omega\in\bar{\omega}_i})$$

by E_i or $E_i(\bar{\omega}_i)$. Define

$$E(\bar{\omega}) = E(x(\bar{\omega}), y(\bar{\omega})) = \cap_{i\in[n]} E_i(x(\bar{\omega}_i)). \tag{B.9}$$

We also write $E(\bar{\omega})$ to denote E. Now, we will prove that given $E(x(\bar{\omega}), y(\bar{\omega}))$ and the other's strategies, proposing a solution that maximizes x will be approximately best-response. Let's restrict our attention to the case in which E happens given $\bar{\omega}$. Suppose that agent $j \neq i$ proposes a candidate to maximize $x(\cdot)$, *i.e.,* $\sigma_j = \operatorname{argmax}_{\omega\in\bar{\omega}_j} x(\omega)$. Then, conditioned on E, agent i's expected payoff of proposing $\omega \in \bar{\omega}_i$ can be obtained as

$$\begin{aligned} u_i(\omega, \sigma_{-i}|E(\bar{\omega})) &= \Pr\left[\,\omega \text{ is winner} \mid E(\bar{\omega})\,\right] \cdot y(\omega) \\ &= \Pr\left[\,x(\omega) \geq x(\omega'), \forall\omega' \in \bar{\omega}_{-i} \mid E(\bar{\omega})\,\right] \cdot y(\omega) \\ &= x(\omega)^{k(n-1)}y(\omega). \end{aligned}$$

Since we conditioned on $E(\bar{\omega})$, for agent i, the solution that maximizes $x(\omega)^{k(n-1)}y(\omega)$ is exactly the

same to the solution that maximizes $x(\omega)$. This verifies that conditioned on $E(\bar{\omega})$ and assuming that the other agent $j \neq i$ proposes a candidate that maximizes $x(\cdot)$, also maximizing $x(\cdot)$ is agent i's best response. Let σ_i^x be a strategy to propose a solution that maximizes $x(\cdot)$ for agent i.

Given that all the other agents j playing σ_j^x for $j \neq i$, we can further obtain the following regarding agent i's utility

$$\begin{aligned}
u_i(\sigma_i^x, \sigma_{-i}) &= \mathbb{E}\left[\, u_i(\sigma_i^x, \sigma_{-i}) \mid E \,\right]\Pr\left[\, E \,\right] + \mathbb{E}\left[\, u_i(\sigma_i^x, \sigma_{-i}) \mid E^c \,\right]\Pr\left[\, E^c \,\right] \\
&\geq \mathbb{E}\left[\, u_i(\sigma_i^x, \sigma_{-i}) \mid E \,\right]\Pr\left[\, E \,\right] \\
&\geq \mathbb{E}\left[\, u_i(\sigma_i', \sigma_{-i}) \mid E \,\right]\Pr\left[\, E \,\right] \\
&= \mathbb{E}\left[\, u_i(\sigma_i', \sigma_{-i}) \,\right] - \mathbb{E}\left[\, u_i(\sigma_i', \sigma_{-i} | E^c) \,\right]\Pr\left[\, E^c \,\right] \\
&\geq \mathbb{E}\left[\, u_i(\sigma_i', \sigma_{-i}) \,\right] - \Pr\left[\, E^c \,\right],
\end{aligned}$$

where the second inequality follows from the fact that given E, playing σ_i^x is weakly dominant over any other strategy σ_i' for agent i. This implies that if we characterize a good lower bound α such that $\Pr\left[\, E \,\right] \geq \alpha$, we have

$$u_i(\sigma_i^x, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i}) - 1 + \alpha,$$

which implies that σ_i^x is $(1-\alpha)$-approximate BNE.

Analyzing the lower bound on $\Pr\left[\, E \,\right]$. Now, we compute the lower bound on $\Pr\left[\, E \,\right]$. Due to the

independence of the utility distributions, we have

$$\Pr\left[\, E(\bar{\omega}) \,\right] = \prod_{i \in [n]} \Pr\left[\, E_i(\bar{\omega}_i) \,\right]. \tag{B.10}$$

We start with the following claim.

Claim B.6.7. $\Pr\left[\, E_i(\bar{\omega}_i) \,\right] = \mathbb{E}\left[\, \prod_{j=2}^{k} \left(1 - \frac{A_{i,j}}{2}\right) \,\right]$, *where* $A_{i,j} = \left(\frac{X_{i,(j)}}{X_{i,(1)}}\right)^{k(n-1)}$ *for* $j \in [k]$.

Proof of the claim. Note that each probability can be computed as the following.

$$\Pr\left[\, E_i(\bar{\omega}_i) \,\right] = \int_{x(\bar{\omega}_i)} \Pr\left[\, E_i(\bar{\omega}_i) \mid x(\bar{\omega}_i) \,\right] f_x(x(\bar{\omega}_i)) dx(\bar{\omega}_i).$$

Given $x(\bar{\omega}_i)$, we can further compute

$$\begin{aligned}
\Pr\left[\, E_i(\bar{\omega}_i) \mid x(\bar{\omega}_i) \,\right] &= \int_{y(\bar{\omega}_i)} \Pr\left[\, E_i(x(\bar{\omega}_i)) \mid x(\bar{\omega}_i), y(\bar{\omega}_i) \,\right] f_y(y(\bar{\omega}_i)) dy(\bar{\omega}_i) \\
&= \Pr\left[\frac{y(\omega_i^*)}{y(\omega)} \geq \left(\frac{x(\omega)}{x(\omega_i^*)}\right)^{k(n-1)}, \forall \omega \in \bar{\omega}_i \mid x(\bar{\omega}_i) \right] \\
&= \Pr\left[\frac{Y_{i,(1)}}{Y_{i,(j)}} \geq \left(\frac{X_{i,(j)}}{X_{i,(1)}}\right)^{k(n-1)}, \forall j \in [k] \mid X_{i,j}, \forall j \in [k] \right].
\end{aligned}$$

Note that this probability is 1 when $j = 1$. Conditioned on the event $Y_{i,(1)} = t$, this probability can be further expanded as

$$\Pr\left[t\left(\frac{X_{i,(1)}}{X_{i,(j)}}\right)^{k(n-1)} \geq Y_{i,(j)} \mid X_{i,j}, \forall j \in [k], Y_{i,(1)} = t \right] = \min\left(1, t\left(\frac{1}{A_{i,j}}\right)\right).$$

Since each solution is i.i.d. sample, we have

$$\Pr\left[\frac{Y_{i,(1)}}{Y_{i,(j)}} \geq \frac{X_{i,(j)}}{X_{i,(1)}}, \forall j \in [k] \mid X_{i,j}, \forall j \in [k]\right] = \prod_{j=2}^{k}\left(\int_{t=0}^{A_{i,j}} t\left(\frac{1}{A_{i,j}}\right)dt + \int_{t=A_{i,j}}^{1} 1dt\right)$$
$$= \prod_{j=2}^{k}\left(1 - \frac{A_{i,j}}{2}\right) \tag{B.11}$$

Hence we have

$$\begin{aligned}
\Pr\left[\, E_i(\bar{\omega}_i)\,\right] &= \Pr\left[\frac{Y_{i,(1)}}{Y_{i,(j)}} \geq \frac{X_{i,(j)}}{X_{i,(1)}}, \forall j \in [k]\right] \\
&= \Pr\left[\frac{Y_{i,(1)}}{Y_{i,(j)}} \geq \frac{X_{i,(j)}}{X_{i,(1)}}, \forall j \in [k] \mid X_{i,j}, \forall j \in [k]\right]\Pr\left[\, X_{i,j}, \forall j \in [k]\,\right] \\
&= \int_{X_i}\prod_{j=2}^{k}\left(1 - \frac{A_{i,j}}{2}\right)dX_i \\
&= \mathbb{E}\left[\prod_{j=2}^{k}\left(1 - \frac{A_{i,j}}{2}\right)\right]
\end{aligned}$$

a

Now we prove that we can decompose the product inside the expectation as follows.

Claim B.6.8. $\mathbb{E}\left[\prod_{j=2}^{k}\left(1-\frac{A_{i,j}}{2}\right)\right] \geq \int_{t=0}^{1}\prod_{j=2}^{k}\left(\mathbb{E}\left[1-\frac{1}{2}\left(\frac{X_{i,(j)}}{t}\right)^{k(n-1)} \mid X_{i,(1)}=t\right]\right)\Pr\left[\, X_{i,(1)}=t\,\right]dt.$

Proof of the claim. Given $X_{i,(1)} = t$, we note that $1 - \frac{A_{i,j}}{2}$ for $j = 2, \ldots, k$ is nonincreasing function over $(X_{i,(2)}, \ldots, X_{i,(k)})$. Now, we use the following correlation inequality called as *Fortuin-Kasteleyn-Ginibre* (FKG) inequality.

Lemma B.6.9 ([212]). *Let $\nu_1, \ldots, \nu_n$ be a probability measure on $\mathbb{R}$. If functions $f, g : \mathbb{R}^n \mapsto [0,1]$ are monotone increasing functions, then $\mathbb{E}_\nu[f \cdot g] \geq \mathbb{E}_\nu[f] \cdot \mathbb{E}_\nu[g]$, where ν is a product measure over $\nu_1, \ldots, \nu_n$. The same inequality holds when both f, g are monotone decreasing functions.*

For given i, we have $f_j = 1 - \frac{A_{i,j}}{2}$ for $j = 2, 3, \ldots, k$, each of which is a nonincreasing function over $X_{i,(j)}$.[3] Hence, by applying FKG inequality, we can further expand it as follows.

$$\begin{aligned}\mathbb{E}\left[\prod_{j=2}^{k}\left(1-\frac{A_{i,j}}{2}\right)\right] &= \int_{t=0}^{1}\mathbb{E}\left[\prod_{j=2}^{k}\left[1-\frac{1}{2}\left(\frac{X_{i,(j)}}{t}\right)^{k(n-1)}\right] \mid X_{i,(1)}=t\right]\Pr\left[X_{i,(1)}=t\right]dt \\ &\geq \int_{t=0}^{1}\prod_{j=2}^{k}\left(\mathbb{E}\left[1-\frac{1}{2}\left(\frac{X_{i,(j)}}{t}\right)^{k(n-1)} \mid X_{i,(1)}=t\right]\right)\Pr\left[X_{i,(1)}=t\right]dt,\end{aligned} \tag{B.12}$$

and it completes the proof. a

Now we show that we can derive a constant lower bound that depends only on k, n and j on the quantity inside the expectation as follows.

Claim B.6.10. *For $j = 2, 3, \ldots, k$, define $g_j(t)$ as follows.*

$$g_j(t) = \mathbb{E}\left[1-\frac{1}{2}\left(\frac{X_{i,(j)}}{t}\right)^{k(n-1)} \mid X_{i,(1)}=t\right].$$

Then, we have the following lower bound on $g_j(t)$.

$$g_j(t) \geq 1-\frac{1}{2}\left(\frac{k-1}{kn-1}\right)^{j-1}.$$

Proof of the claim. We remark that $g_j(t)$ is a lower bound on the probability that $E_i(\bar{\omega}_i)$ happens

[3]Note in fact that we cannot directly plug in $f_j = 1 - A_{i,j}/2$ in the inequality due to the correlation between the order statistics. That said, by rewriting the expression to be over the original random variables instead of order statistics, we can easily see that the same result follows, since the expectation remains for either of expressions.

conditioned on $X_{i,(1)} = t$. Rewrite $g_j(t)$ as

$$\int_{x=0}^{t} \left(1 - \frac{1}{2}\left(\frac{x}{t}\right)^{k(n-1)}\right) f_{X_{i,(j)}|X_{i,(1)}=t}(x)dx$$

Now, the joint density function of $X_{i,(j)}, X_{i,(1)}$ can be computed as

$$\begin{aligned} f_{X_{i,(j)},X_{i,(1)}}(x,t) &= f_{U_{(k-j+1)},U_{(k)}}(x,t) \\ &= k!\frac{x^{k-j}}{(k-j)!}\frac{(t-x)^{j-2}}{(j-2)!}, \end{aligned}$$

where $U_{(j)}$ denotes j-th order statistics from k i.i.d. random variables with $U[0,1]$.

The density function of $X_{i,(1)}$ is p.d.f. of $Beta(k,1)$, which is $t^{k-1}/B(k,1)$ where $B(\cdot,\cdot)$ is the beta function. This implies that

$$f_{X_{i,(j)}|X_{i,(1)}=t}(x) = k!\frac{x^{k-j}}{(k-j)!}\frac{(t-x)^{j-2}}{(j-2)!}\frac{B(k,1)}{t^{k-1}}$$

Hence, we have

$$g_j(t) = \int_{x=0}^{t} \left(1 - \frac{1}{2}\left(\frac{x}{t}\right)^{k(n-1)}\right) \frac{k!}{(k-j)!(j-2)!}\frac{B(k,1)}{t^{k-1}} x^{k-j}(t-x)^{j-2}dx.$$

By change of variables with $x = ty$, and we have

$$\begin{aligned} g_j(t) &= \int_{y=0}^{1} \left(1 - \frac{y^{k(n-1)}}{2}\right) \frac{k!}{(k-j)!(j-2)!}\frac{B(k,1)}{t^{k-1}} (ty)^{k-j}(t-ty)^{j-2}tdy \\ &= \frac{k!B(k,1)}{(k-j)!(j-2)!}\int_{y=0}^{1} \left(1 - \frac{y^{k(n-1)}}{2}\right) y^{k-j}(1-y)^{j-2}dy, \end{aligned}$$

where we use the fact that $B(x,y) = \frac{(x-1)!(y-1)!}{(x+y-1)!}$ for positive integers x, y, Rewriting the RHS using the definition of beta function, we can further expand

$$g_j(t) = \frac{(k-1)!}{(k-j)!(j-2)!}\left(B(k-j+1,j-1) - \frac{B(kn-j+1,j-1)}{2}\right) = 1 - \frac{1}{2}\frac{(kn-j)!(k-1)!}{(kn-1)!(k-j)!}$$
$$= 1 - \frac{1}{2}\frac{(k-j+1)\dots(k-1)}{(kn-j+1)\dots(kn-1)}$$
$$\geq 1 - \frac{1}{2}\left(\frac{k-1}{kn-1}\right)^{j-1}, \tag{B.13}$$

where we use that $B(x,y) = \int_{t=0}^{1} t^{x-1}(1-t)^{y-1}dt$ for any real x, y. For the inequality, we use $\frac{k-j+1}{kn-j+1} \leq \frac{k-1}{kn-1}$ for any $j = 2, 3, \dots, k$. This completes the proof. a

Due to the claims above (Claim B.6.7,B.6.8 and B.6.10) and by (B.10), we obtain the following lower bound on the event E_i.

$$\begin{aligned}
\Pr\left[E_i\right] &= \mathbb{E}\left[\prod_{j=2}^{k}\left(1 - \frac{A_{i,j}}{2}\right)\right] \\
&\overset{(a)}{\geq} \int_{t=0}^{1}\left(\prod_{j=2}^{k} g_j(t)\right)\Pr\left[X_{i,(1)} = t\right]dt \\
&\overset{(b)}{\geq} \int_{t=0}^{1}\left(\prod_{j=2}^{k} 1 - \frac{1}{2}\left(\frac{k-1}{kn-1}\right)^{j-1}\right)\Pr\left[X_{i,(1)} = t\right]dt \\
&= \prod_{j=2}^{k}\left(1 - \frac{1}{2}\left(\frac{k-1}{kn-1}\right)^{j-1}\right) \\
&\geq \prod_{j=2}^{k}\left(1 - \frac{1}{2}\left(\frac{1}{n}\right)^{j-1}\right),
\end{aligned}$$

where (a) holds by Claim B.6.8 and (b) follows from Claim B.6.10.

Using $1 - x \geq e^{-x/(1-x)}$ for $x \in (0, 1)$, we finally have

$$\Pr\left[\, E(\bar{\omega}_i) \,\right] \geq \exp\left(-\sum_{j=2}^{k} \frac{\left(\frac{1}{n}\right)^{j-1}}{2 - \left(\frac{1}{n}\right)^{j-1}} \right). \tag{B.14}$$

Now we proceed with the following series of algebraic manipulation.

$$\begin{aligned}
(B.14) &\geq \exp\left(-\sum_{j=2}^{k} \frac{1}{2n^{j-1} - 1} \right) \\
&\geq \exp\left(-\sum_{j=2}^{k} \frac{1}{2} \frac{1}{n^{j-1} - 1} \right) \\
&= \exp\left(-\frac{1}{2(n-1)} \sum_{j=2}^{k} \frac{1}{n^{j-2} + n^{j-3} + \ldots + 1} \right) \\
&\geq \exp\left(-\frac{1}{2(n-1)} \sum_{j=2}^{k} \frac{1}{n^{j-2}} \right) \\
&= \exp\left(-\frac{1}{2(n-1)} \frac{1 - \left(\frac{1}{n}\right)^{k-1}}{1 - \frac{1}{n}} \right) \\
&\geq \exp\left(-\frac{n}{2(n-1)^2} \right).
\end{aligned}$$

Multiplying it over $i \in [n]$, we have

$$\Pr\left[\, E \,\right] \geq \exp\left(-\frac{n^2}{2(n-1)^2} \right), \tag{B.15}$$

where it converges to $1/\sqrt{e} \approx 0.6065...$ as n increases. Hence, we conclude that playing σ_i^x is $(1 - e^{-n^2/2(n-1)^2})$-approximate BNE, and as the principal's utility is $X_{\max}$ for any observation of solutions under this equilibrium, we complete the proof. a

B.6.16 Proof of Theorem 3.7.5

Proof. Due to our definition of RSPM, B of proposed candidates will be randomly chosen by the principal and the best candidate among them will be selected as the winner. Let E be the set of agents who have at least one eligible solution to propose, and E' be the set of other agents. Denote the size of sets E and E' by e and e' respectively so that $e + e' = n$. Let $\bar{\omega} = \cup_{i\in[n]}\bar{\omega}_i$ be the entire set of solutions observed by the agents and $\sigma = (\sigma_1, \ldots, \sigma_n)$ be arbitrary equilibrium strategy. For $i \in E'$, since agent i does not have eligible solutions, even his best-observed solution which gives the utility of $X_{i,(1)}$ to the principal is not eligible. This implies that $X_{i,(1)} < \tau$. On the other hand, for $i \in E$, agent i has at least one eligible observed solution to propose which implies that his best-observed solution is also eligible as well, *i.e.,* $X_{i,(1)} \geq \tau$. These arguments further imply that $X_{(e)} \geq \tau > X_{(e+1)}$. We consider two cases (i) when $e \leq B$ and (ii) when $e > B$, and prove the theorem separately.

Case (i): $e \leq B$. In this case, the number of agents who have eligible solutions is at most B, which implies that $X_{(B+1)} < \tau$. In this case, there exist at most B proposed candidates thus the principal can examine all of them and take the best one. As a result, this case is equivalent to the unlimited budget we had in Theorem 3.5.1, and hence from the same argument, the principal's expected utility is at least

$$\mathbb{E}\left[X_{(2)}1\left[X_{(2)} \geq \tau > X_{(B+1)}\right]\right] + \mathbb{E}\left[\tau 1\left[X_{(1)} \geq \tau > X_{(2)}\right]\right].$$

Case (ii): $e > B$. We follow the proof steps of Theorem 3.5.1. Assume that agents play mixed strategies, *i.e.,* they construct a probability distribution over the observed solutions. Formally, given the solutions $\bar{\omega}$, denote by $p_{i,j}$, the probability that agent i submits $\omega_{i,j}$ as his candidate. These probability

distributions shape a mixed Nash equilibrium.

For each $i \in [n]$ we have $\sum_{j=1}^{k_i} p_{i,j} = 1$. Define $S_i = \{\omega_{i,j} : p_{i,j} > 0\}$ as the multi-set of solutions that agent i proposes with positive probability. For each agent i, let b_i be the candidate which gives the worst utility to the principal, *i.e.,* $b_i = \operatorname{argmin}_{\omega \in S_i} x(\omega)$. If there are multiple candidates with the worst utility, we arbitrarily choose one. We use A to denote the set of agents who have non-zero utility.

We proceed with the following claim.

Claim B.6.11. $|A| \leq e - B + 1$.

Proof of the claim. Suppose not, then $|A| \geq e - B + 2$. Let $t \in A$ be an agent such that for each agent $i \in A \setminus \{t\}$, either

- $x(b_t) < x(b_i)$
- $x(b_t) = x(b_i)$, and mechanism prefers agent i over agent t using tie-breaking rule ρ,

i.e., agent t is the agent with non-zero utility having worst b_t and least preferred using ρ. As he has non-zero utility and $b_t \in S_i$, proposing b_t gives him non-zero utility as well.

According to the definition of b_t, when t proposes b_t, no matter what agents in $A \setminus \{t\}$ propose, their candidates are dominant, *i.e.,* the mechanism prefers candidate submitted by any agent in $A \setminus \{t\}$ over b_t.

There are e candidates in total and as $|A \setminus \{t\}| \geq e - B + 1$, any random choice of B candidates includes at least one candidate of agents in $A \setminus \{t\}$. This implies that b_t is never selected as the winner, which is a contradiction with the fact that proposing b_t gives non-zero utility to agent t. So, we conclude that $|A| \leq e - B + 1$. a

Now, we focus on agents in $E \setminus A$. Let $t' \in E \setminus A$ be an agent such that for any agent $j \in E \setminus (A \cup \{t'\})$, either

- $X_{t',(1)} > X_{j,(1)}$
- $X_{t',(1)} = X_{j,(1)}$, and mechanism prefers agent t' over agent j using tie-breaking rule ρ,

i.e., agent t' has the first-best solution among agents in $E \setminus A$ and mostly preferred by tie-breaking rule ρ. Note that we proved $|A| \leq e - B + 1$ so $|E \setminus A| \geq B - 1$. Now we claim $X_{t',(1)} \leq x(b_t)$ as follows.

Claim B.6.12. $X_{t',(1)} \leq x(b_t)$.

Proof of the claim. Suppose not and let $x(b_t) < X_{t',(1)}$. We know that under Nash equilibrium, agent t' gets zero utility thus given the strategies of others, no matter what his strategy is, he gets zero utility. As a result, even if he proposes his first-best solution that gives the utility of $X_{t',(1)}$ to the principal, his solution is never selected. However, according to the definition of t', this solution is always preferred to all candidates proposed by agents in $E \setminus (A \cup \{t'\})$ which are at least $B - 2$ candidates. It is also dominant over b_t. Hence due to the randomness of the mechanism, there exists a non-zero probability that the principal takes these $B - 1$ candidates and the candidate proposed by t' so that t' becomes the winner and gets non-zero utility. This is a contradiction to the fact that t' has zero utility and thus $X_{t',(1)} \leq x(b_t)$. a

Let ω^* be the winner. ω^* belongs to agents in A and according to definition of b_t, $x(b_t) \leq x(\omega^*)$. We also proved that $X_{t',(1)} \leq x(b_t)$ so principal's utility is at least $X_{t',(1)}$. We recall that

- According to definitions of E and E', $X_{t',(1)} \geq \tau > X_{i,(1)}$ for all $i \in E'$.
- According to the definition of t', we have $X_{t',(1)} \geq X_{j,(1)}$ for each $j \in E \setminus (A \cup \{t'\})$.

Hence $X_{t',(1)}$ is at least as large as the first-best solutions of $|E'|+|E\setminus(A\cup\{t'\})\,|\geq n-e+B-2$ agents. Therefore, we can conclude that $X_{t',(1)}\geq X_{(e-B+2)}$. So, in this case where $X_{(e)}\geq\tau>X_{(e+1)}$, we obtain a lower bound on principal's utility as

$$\mathbb{E}\left[\,X_{(e-B+2)}1\left[X_{(e)}\geq\tau>X_{(e+1)}\right]\,\right].$$

Now by iterating over all values for $e>B$, we state that the expected utility is at least

$$\sum_{e=B+1}^{n}\mathbb{E}\left[\,X_{(e-B+2)}1\left[X_{(e)}\geq\tau>X_{(e+1)}\right]\,\right],$$

where we write $X_{(n+1)}:=-\infty$. By combining the result of case (i), and case (ii), we can infer that the principal's expected utility is at least

$$\mathbb{E}\left[\,X_{(2)}1\left[X_{(2)}\geq\tau>X_{(B+1)}\right]\,\right]+\mathbb{E}\left[\,\tau 1\left[X_{(1)}\geq\tau>X_{(2)}\right]\,\right]+\sum_{e=B+1}^{n}\mathbb{E}\left[\,X_{(e-B+2)}1\left[X_{(e)}\geq\tau>X_{(e+1)}\right]\,\right],$$

which shows that the additive approximation factor is at most

$$\begin{aligned}\mathbb{E}\left[\,X_{(1)}\,\right]-\mathbb{E}\left[\,X_{(2)}1\left[X_{(2)}\geq\tau>X_{(B+1)}\right]\,\right]-\mathbb{E}\left[\,\tau 1\left[X_{(1)}\geq\tau>X_{(2)}\right]\,\right]\\-\sum_{e=B+1}^{n}\mathbb{E}\left[\,X_{(e-B+2)}1\left[X_{(e)}\geq\tau>X_{(e+1)}\right]\,\right]\end{aligned}$$

which finishes the proof.

a

B.6.17 Proof of Theorem B.3.2

Proof. We recommend the readers skim the proof of Theorem 3.6.7 in Appendix B.6.15 before reading this proof. Similar to the proof of Theorem 3.6.7, we proceed with three main steps. First, we construct a Bayes correlated equilibrium with a certain choice of signaling scheme in which we signal the agents whenever some good event E happens. Next, we show that the principal's utility given the signal can be decomposed as $\mathbb{E}[X_{\max}] \cdot \Pr[E]$. Finally, we obtain a lower bound on the probability that E happens.

Constructing Bayes correlated equilibrium. Consider an MSPM with no eligible sets and a deterministic tie-breaking rule ρ. Denote by $\bar{\omega} = \cup_{i\in[n]}\bar{\omega}_i$ the solutions observed by the agents. Similar to the proof of Theorem 3.6.7, define an event $E_i(\cdot)$ as follows.

$$E_i = \left\{ x(\bar{\omega}_i), y(\bar{\omega}_i) : \operatorname*{argmax}_{\omega\in\bar{\omega}_i} x(\omega)^{k(n-1)} y(\omega) = \operatorname*{argmax}_{\omega\in\bar{\omega}_i} x(\omega) \right\}.$$

Let $x(A) = \{x(\omega)\}_{\omega\in A}$ and similarly for $y(A)$. Define $E = \cap_{i\in[n]} E_i$. Let σ_i^x be a strategy to propose a solution that maximizes $x(\cdot)$ for agent i. Denote by $u_i(\sigma_i, \sigma_{-i}|A)$ agent i's expected utility conditioned on the event A, *i.e.,*

$$u_i(\sigma_i, \sigma_{-i}|A) = \mathbb{E}\left[\int_{\bar{\omega}\in A} y_i(f_{M,(\sigma_i,\sigma_{-i})}) f(\bar{\omega}|A) d\bar{\omega} \right],$$

where we write $f(\bar{\omega}|A)$ to denote the probability density function of $\bar{\omega}$ conditioned on A. Conditioned

on E_i, agent i's expected payoff of proposing a solution $\omega \in \bar{\omega}_i$ given that others play σ^x_{-i} is

$$\begin{aligned} u_i(\omega, \sigma^x_{-i}|E_i) &= \Pr\left[\, \omega \text{ is winner} \mid E \,\right] \cdot y(\omega) \\ &= \Pr\left[\, x(\omega) \geq x(\omega'), \forall \omega' \in \bar{\omega}_{-i} \mid E \,\right] \cdot y(\omega) \\ &= x(\omega)^{k(n-1)} y(\omega). \end{aligned}$$

This implies that conditioned on E_i, agent i's best response given that others play σ^x_{-i} is also σ^x_i. Hence, conditioned on $E(\bar{\omega})$, playing $\sigma^x = (\sigma^x_1, \ldots, \sigma^x_n)$ is best-response for each agent given the others' strategies.

Observe that due to the Nash's existence theorem, there exists at least one BNE. Denote this equilibrium by $\sigma^o = (\sigma^o_1, \ldots, \sigma^o_n)$. Now consider a mediator who recommends $\psi_i = \sigma^x_i$ to agent $i \in [n]$ whenever $E(\bar{\omega})$ happens, and recommends σ'_i otherwise. Due to our construction of E, we have

$$u_i(\psi_i, \sigma^x_{-i}|E) \geq u_i(\sigma_i, \sigma^x_{-i}|E),$$

for any other strategy σ'_i of agent i. From our definition of σ^o, we also have

$$u_i(\psi_i, \sigma^o_{-i}|\neg E) \geq u_i(\sigma'_i, \sigma^o_{-i}|\neg E),$$

for any other strategy σ'_i of agent i. Combining two inequalities above, we can verify that obeying

$\psi = (\psi_1, \ldots, \psi_n)$ constructs a BNE as follows

$$\begin{aligned} u_i(\psi_i, \psi_{-i}) &= u_i(\psi_i, \sigma^x_{-i}|E)\Pr[\,E\,] + u_i(\psi_i, \sigma^o_{-i}|\neg E)\Pr[\,\neg E\,] \\ &= u_i(\psi_i, \sigma^x_{-i}|E)\Pr[\,E\,] + u_i(\psi_i, \sigma^o_{-i}|\neg E)\Pr[\,\neg E\,] \\ &\geq u_i(\sigma'_i, \sigma^x_{-i}|E)\Pr[\,E\,] + u_i(\sigma'_i, \sigma^o_{-i}|\neg E)\Pr[\,\neg E\,]. \end{aligned}$$

Decomposing the principal's utility. For notational simplicity, denote by $x(\sigma)$ the principal's expected utility given the strategies σ, and $x(\sigma|\bar{\omega})$ the principal's utility given $\bar{\omega}$, *i.e.,* given all the random variables $x(\bar{\omega})$ and $y(\bar{\omega})$. Denote by $f_x(x(\bar{\omega}))$ and $f_y(y(\bar{\omega}))$ the p.d.f. of $x(\bar{\omega})$ and $y(\bar{\omega})$, respectively. Now, under the BCE with ψ, suppose that the agents obey the recommended strategies, *i.e.,* $\sigma = \psi$. The principal's expected utility $x(\sigma)$ can be lower bounded as

$$\begin{aligned} &\int_{x(\bar{\omega})} \left(\int_{y(\bar{\omega})} \mathbb{E}\,[\,x(\sigma|\bar{\omega}) \mid x(\bar{\omega}), y(\bar{\omega})\,] f_y(y(\bar{\omega}))dy(\bar{\omega}) \right) f_x(x(\bar{\omega}))dx(\bar{\omega}) \\ &\overset{(a)}{=} \int_{x(\bar{\omega})} \left(\int_{y(\bar{\omega})} \mathbb{E}\,[\,x(\sigma|\bar{\omega}) \mid x(\bar{\omega}), y(\bar{\omega}), E\,]\Pr[\,E \mid x(\bar{\omega}), y(\bar{\omega})\,] f_y(y(\bar{\omega}))dy(\bar{\omega}) \right) \\ &+ \left(\int_{y(\bar{\omega})} \mathbb{E}\,[\,x(\sigma|\bar{\omega}) \mid x(\bar{\omega}), y(\bar{\omega}), \neg E\,]\Pr[\,\neg E \mid x(\bar{\omega}), y(\bar{\omega})\,] f_y(y(\bar{\omega}))dy(\bar{\omega}) \right) f_x(x(\bar{\omega}))dx(\bar{\omega}) \\ &\overset{(b)}{\geq} \int_{x(\bar{\omega})} \left(\int_{y(\bar{\omega})} \mathbb{E}\,[\,x(\sigma|\bar{\omega}) \mid x(\bar{\omega}), y(\bar{\omega}), E\,]\Pr[\,E \mid x(\bar{\omega}), y(\bar{\omega})\,] f_y(y(\bar{\omega}))dy(\bar{\omega}) \right) f_x(x(\bar{\omega}))dx(\bar{\omega}), \end{aligned}$$

where (a) holds by the law of total expectation, and (b) follows from ignoring the latter integral, *i.e.,* the one conditioned on $\neg E$. First, we observe that conditioned on the event E, the principal's utility $x(\sigma|\bar{\omega})$ is simply $X_{\max}$. This is because conditioned on E, each agent i submits σ^x_i, and the principal finally chooses the one that maximizes her own utility. By the independence of $X_{\max}$ over

the random variables $y(\cdot)$, we can further obtain

$$x(\sigma) \geq \int_{x(\bar{\omega})} \left(\int_{y(\bar{\omega})} \Pr\left[E \mid x(\bar{\omega}), y(\bar{\omega}) \right] f_y(y(\bar{\omega})) dy(\bar{\omega}) \right) X_{\max} f_x(x(\bar{\omega})) dx(\bar{\omega}). \tag{B.16}$$

$$= \int_{x(\bar{\omega})} \prod_{i=1}^{n} \left(\int_{y_i(\bar{\omega}_i)} \Pr\left[E \mid x(\bar{\omega}), y_i(\bar{\omega}) \right] f_{y_i}(y_i(\bar{\omega})) dy_i(\bar{\omega}) \right) X_{\max} f_x(x(\bar{\omega})) dx(\bar{\omega}), \tag{B.17}$$

where the last equality follows from the independence of the solutions between the agents. Defining $A_{i,j} = \left(\frac{X_{i,(j)}}{X_{i,(1)}} \right)^{k(n-1)}$, due to (B.11), we can rewrite (B.17) as

$$\int_{x(\bar{\omega})} \left(\prod_{i=1}^{n} \prod_{j=2}^{k} \left(1 - \frac{A_{i,j}}{2} \right) \right) X_{\max} f_x(x(\bar{\omega})) dx(\bar{\omega}) \tag{B.18}$$

$$= \mathbb{E}\left[X_{\max} \prod_{i=1}^{n} \prod_{j=2}^{k} \left(1 - \frac{A_{i,j}}{2} \right) \right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[t \prod_{i=1}^{n} \prod_{j=2}^{k} \left(1 - \frac{A_{i,j}}{2} \right) \mid X_{\max} = t \right] \right]. \tag{B.19}$$

$$= \int_{t=0}^{1} \mathbb{E}\left[\prod_{i=1}^{n} \prod_{j=2}^{k} \left(1 - \frac{A_{i,j}}{2} \right) \mid X_{\max} = t \right] t \Pr\left[X_{\max} = t \right] dt \tag{B.20}$$

Now we prove the following claim.

Claim B.6.13. *$g(t) = \mathbb{E}\left[\prod_{i=1}^{n} \prod_{j=2}^{k} \left(1 - \frac{A_{i,j}}{2} \right) \mid X_{\max} = t \right]$ is an increasing function on t.*

Proof of the claim. We can write $g(t)$ as

$$\sum_{s=1}^{n} \mathbb{E}\left[\prod_{i=1}^{n} \prod_{j=2}^{k} \left(1 - \frac{A_{i,j}}{2} \right) \mid X_{s,(1)} = t, \operatorname*{argmax}_{i \in [n]} X_{i,(1)} = s \right] \Pr\left[\operatorname*{argmax}_{i \in [n]} X_{i,(1)} = s \mid X_{\max} = t \right].$$

Due to symmetry, we know $\Pr\left[\operatorname{argmax}_{i \in [n]} X_{i,(1)} = s \mid X_{\max} = t \right]$ is exactly equal for $s \in [n]$, and it does not change with respect to t. In addition, given $X_{s,(1)} = t$ and $\operatorname{argmax}_{i \in [n]} X_{i,(1)} = s$, in

$\prod_{i=1}^{n}\prod_{j=2}^{k}\left(1-\frac{A_{i,j}}{2}\right)$, only the terms $\prod_{j=2}^{k}\left(1-\frac{A_{s,j}}{2}\right)$ changes with respect to t. Given $X_{s,(1)}=t$, we can rewrite it as

$$\prod_{j=2}^{k}\left(1-\frac{A_{s,j}}{2}\right)=\prod_{j=2}^{k}\left(1-\left(\frac{X_{s,(j)}}{t}\right)^{k(n-1)}\right),$$

and it is obviously an increasing function with respect to t. Combining the results, we conclude that $g(t)$ is an increasing function on t. a

Rewriting the expectation based on $X_{i,j}$ and applying FKG inequality similar to the proof of Theorem 3.6.7, we obtain

$$\begin{aligned}(B.20) &\geq \left(\int_{t=0}^{1}\mathbb{E}\left[\prod_{i=1}^{n}\prod_{j=2}^{k}\left(1-\frac{A_{i,j}}{2}\right)\mid X_{\max}=t\right]\Pr\left[X_{\max}=t\right]t\right)\cdot\left(\int_{t=0}^{1}t\Pr\left[X_{\max}=t\right]dt\right)\\ &=\mathbb{E}\left[\prod_{i=1}^{n}\prod_{j=2}^{k}\left(1-\frac{A_{i,j}}{2}\right)\right]\mathbb{E}\left[X_{\max}\right]\\ &=\Pr\left[E\right]\cdot\mathbb{E}\left[X_{\max}\right].\end{aligned}$$

Obtaining a lower bound on $\Pr\left[E\right]$**.** Now it remains to lower bound the probability $\Pr\left[E\right]$, which is already given in (B.15). Hence, we conclude that the principal's expected utility under this BCE is at least

$$\exp\left(-\frac{n^2}{2(n-1)^2}\right)\mathbb{E}\left[X_{\max}\right],$$

and it completes the proof. a

B.6.18 Proof of Theorem B.5.3

Proof. We start by presenting the following lemma from [3].

Lemma B.6.14 ([3]). *In the single-agent setting, for any mechanism M and corresponding dominant strategies σ, there exists a MSPM such that for any equilibrium strategies σ', we have $x(f_{M,\sigma}(\bar{\omega})) = x(f_{SPM,\sigma'}(\bar{\omega}))$ for any $\bar{\omega} \in \Omega^*$.*

By the lemma, it suffices to consider the case when M is a SPM. Let the realized solutions in P be $\bar{\omega} = \{\omega_1, \omega_2, \ldots, \ldots \omega_k\}$. Consider a partition of $\bar{\omega}$ as follows

$$\bar{\omega}_1 = \{\omega_1, \ldots, \omega_{k_1}\}$$

$$\bar{\omega}_2 = \{\omega_{k_1+1}, \ldots, \omega_{k_1+k_2}\}$$

$$\vdots$$

$$\bar{\omega}_n = \{\omega_{k_1+\ldots+k_{n-1}+1}, \ldots, \omega_{k_1+\ldots+k_n=k}\}.$$

Suppose that agent i's realization of solutions in P' is $\bar{\omega}_i$ so that the outcomes of P and P' are coupled. This coupling is indeed possible because the probability that agent i's solutions are realized to be $\bar{\omega}_i = \{\omega_{i_1}, \ldots, \omega_{i_{k_i}}\}$ is exactly equal to the marginal probability that ω's i-th partition is realized to be $\bar{\omega}_i = \{\omega_{i_1}, \ldots, \omega_{i_{k_i}}\}$.

Consider a SPM with an eligible set R under P, and corresponding equilibrium strategies σ. Denote by x and y the principal and the agent's utility function under P respectively. Given the solutions realized as above, SPM will select the solution which maximizes $y(\omega_j)$ for $j \in [k]$ such that $\omega_j \in R$, *i.e.,* $\omega_1^* = \operatorname{argmax}_{j\in[n],\omega_j\in R} y(\omega_j)$. Let this solution be ω_1^*. Now consider a MSPM

with arbitrary tie-breaking rule and the same eligible set R under P', and suppose that the agents play equilibrium strategies σ'. Let ω_2^* be the corresponding winner in this case.

We want to show that $x(\omega_2^*) \geq x(\omega_1^*)$ for any realization of solutions $\bar{\omega}$. Suppose not. It is obvious that both $x(\omega_1^*)$ and $x(\omega_2^*)$ belong to the eligible set R. If ω_1^* and ω_2^* belongs to the same partition $\bar{\omega}_i$, it means that agent i has both the solutions ω_1^* and ω_2^*. By our definition on ω_1^* such that $\omega_1^* = \text{argmax}_{j\in[n],\omega_j\in R}\, y(\omega_j)$, submitting ω_1^* is gives the better or the same utility for agent i. If $y(\omega_1^*) > y(\omega_2^*)$, then simply submitting ω_1^* gives a strictly better utility for agent i, and it contradicts that σ' is equilibrium strategy. If $y(\omega_1^*) = y(\omega_2^*)$, then $x(\omega_1^*) \leq x(\omega_2^*)$ by Assumption 3.4.9. Hence, $x(\omega_1^*) \leq x(\omega_2^*)$ in both cases.

Now suppose that ω_1^* and ω_2^* belongs to the different partitions $\bar{\omega}_i$ and $\bar{\omega}_j$. Suppose that agent i's equilibrium strategy satisfies $\sigma_i' \neq \omega_1^*$. If playing $\bar{\omega}_i$ is strictly better for agent i, it contradicts that σ_i' is an equilibrium strategy. Hence $y(\omega_1^*) = y(\sigma_i')$. Again by Assumption 3.4.9, since agent i plays σ_i', we have $x(\sigma_i') \geq x(\omega_1^*)$. Note that the principal should observe both σ_i' and ω_2^*, and commits to ω_2^*. This implies that $x(\omega_2^*) \geq x(\sigma_i') \geq x(\omega_1^*)$, and we finish the proof. a

Appendix C: Replication-proof Bandit Mechanism Design with Bayesian Agents

C.1 Related Works

There exists a growing literature at the intersection of multi-armed bandits and algorithmic game theory. The work of [111] considers a setting where the arms are agents. When an arm (agent) is pulled a reward is realized. The agent can take part of the reward and only report the remaining reward to the principal. The work of [112] is closely related to [111] in that the agents are still arms that can modify their rewards. However, in [112] an agent's utility equals the total number of pulls and are not related to the rewards. A number of recent works [12, 213, 214] in algorithmic game theory community study the problem of similar spirits, but the problem setups are very different from ours.

Another growing line of works is in the incentivized exploration, initiated by [215] and [216]. To incentivize a myopic agent arriving at each round and pulling the arm, the principal needs to design an incentive-compatible signaling scheme to induce myopic agents to explore the arms. To this end, the principal may exploit the information asymmetry [217–219] or may compensate the myopic users [220, 221]. Another line of works [222–225] has been asking a similar question of *learning dynamics* under a presence of myopic agents pulling the arms, not in a mechanism design perspective, but to analyze the conseqeunce of the given game.

C.2 Motivations

Despite being directly motivated by problem posed by [22], we here present further motivation in terms of the real world applications. The content replication phenomenon in content platforms like

Youtube [226] or ad replication in ad platforms such as Google AdSense [227] would be good practical applications. In contents platforms like Youtube, a content creator (agent) typically does not exactly know about each content's quality in advance. Historical information on the contents can be thought of as a prior distribution of the agent. The setting by [22] can be interpreted as a scenario in which all the agents know the exact quality of their contents, which is not practical. Our Bayesian setting indeed captures more realistic scenario of each agent having only a partial information on the expected outcome of each content. We finally remark that although the problem of replication-proof mechanism has not studied in the bandit literature except [22], there has been a number of papers for sybil-proof mechanism design in the algorithmic game theory literature [228, 229].

C.3 Extension to the Asymmetric Setting

Our algorithm and the analysis therein heavily relies on the symmetry assumption such that each agent has a single distribution which is shared over all its arms. Indeed, our analysis does not directly extend to the asymmetric case. In particular, in the asymmetric setting in which each arm is possibly equipped with a different prior distribution, each agent should determine specific number of replication for each arm. This corresponds to possibly very-different distribution over "bandit instances", which we believe to be significantly challenging to compare in a theoretical manner. We firmly believe that such an extension would be very interesting direction from both practical and theoretical perspectives. However, one can find some evidences on dominant strategies for certain cases in the asymmetric setting. For example, if an agent has two types of arms each sampling from distributions F_1 and F_2 and if F_1 is stochastically dominated by F_2, then one can easily argue that F_2 is dominant strategy over F_1, so it reduces to the symmetric case with F_2.

C.4 Failure of Existing Algorithms

We here provide more details on the failure of existing algorithms, presented in Theorem 5.3.1. First, the following is the pseudocode of the standard UCB algoritm in the literature.

ALGORITHM 5: UCB [32]

Input: Tie-breaking rule τ, arm set A
Let $\hat{\mu}_a \leftarrow 0$ and $n_a \leftarrow 0$ for $a \in A$
for $t = 1, 2, \ldots$ **do**
  if $n_a = 0$ *for some* $a \in A$ **then**
    Let $a^* \leftarrow a$
  else
    Compute $s_a \leftarrow \hat{\mu}_a + \sqrt{2 \ln t / n_a}$
    Let $\hat{a} \leftarrow \mathrm{argmax}_{a \in A}\, s_a$
Pull $\hat{a}$ and obtain reward R_t
Update statistics: $\hat{\mu}_{\hat{a}} \leftarrow \frac{n_{\hat{a}} \hat{\mu}_{\hat{a}} + R_t}{n_{\hat{a}} + 1}$; $n_{\hat{a}} \leftarrow n_{\hat{a}} + 1$

This is extended to the H-UCB, whose pseudocode is presented as follows.

ALGORITHM 6: H-UCB [22]

Input: Tie-break rule, agent set $\mathcal{N}$, arm set $\mathcal{S}_i$ for $i \in \mathcal{N}$
Let $\hat{\mu}_i \leftarrow 0$, $n_i \leftarrow 0$, $\hat{\mu}_{i,a} \leftarrow 0$, $n_{i,a} \leftarrow 0$ for $i \in \mathcal{N}$ and $a \in \mathcal{S}_i$
for $t = 1, 2, \ldots$ **do**
  if $n_i = 0$ *for some* $i \in \mathcal{N}$ **then**
    Let $\hat{i} \leftarrow i$
  else
    Compute $s_i = \hat{\mu}_i + \sqrt{2 \ln(t) / n_i}$
    Let $\hat{i} \leftarrow \mathrm{argmax}_{i \in \mathcal{N}}\, s_i$
  if $n_{\hat{i},a} = 0$ *for some* $a \in \mathcal{S}_{\hat{i}}$ **then**
    Let $\hat{a} \leftarrow a$
  else
    Compute $s_{\hat{i},a} \leftarrow \hat{\mu}_{\hat{i},a} + \sqrt{2 \ln(n_i) / n_{a,t}}$
  Pull $\hat{a}$ of agent $\hat{i}$ and obtain reward R_t
  Update statistics: $\hat{\mu}_{\hat{i},\hat{a}} \leftarrow \frac{\hat{\mu}_{\hat{i},\hat{a}} n_{\hat{i},\hat{a}} + R_t}{n_{\hat{i},\hat{a}} + 1}$; $\hat{\mu}_{\hat{i}} \leftarrow \frac{\hat{\mu}_{\hat{i}} n_{\hat{i}} + R_t}{n_{\hat{i}} + 1}$; $n_{\hat{i},\hat{a}} \leftarrow n_{\hat{i},\hat{a}} + 1$; $n_{\hat{i}} \leftarrow n_{\hat{i}} + 1$;

As have been noted in the main body of the paper, H-UCB mainly consists of two phases: (i) in the first phase, it runs UCB1 by considering each agent as a single arm and selects one agent at each round, and (ii) in the second phase, it runs UCB within the selected agent. Precisely, it maintains two

types of statistics: (i) agent-wise empirical average reward $\hat{\mu}_i$, number of pulls per agent n_i, and (ii) arm-wise empirical average reward $\hat{\mu}_{i,a}$ and number of pulls per arm $n_{i,a}$.

[22] prove that H-UCB admits replication-proofness while maintaining sublinear regret. Their analysis is largely based on the observation that fully-informed agents *do not* have any incentive to register *suboptimal arms*. Therefore, their setting essentially reduces to the case under which each agent is equipped with only a *single arm*, and its only choice is to determine how many times to replicate the optimal arm. This observation largely simplifies the equilibrium analysis since replicating the same arm does not change the expected rewards of the agent due to the hierarchical structure therein. For example, if one uses uniform-random selection in Phase 1, each agent's expected utility is totally independent from each other. Thus, it suffices to show truthfulness under single-agent setting. Especially, if there is only a single agent, *any standard sublinear regret algorithm* would achieve replication-proofness due to the observation above. Hence, it follows that the pair of uniform-random selection algorithm in the first phase and any standard bandit algorithm in the second phase would achieve replication-proofness.

In our setting, however, the agents only know about the prior distributions from which the arms' mean rewards are drawn, and they are oblivious to which original arm will have the highest mean reward. This motivates each agent to register multiple arms possibly with some replications. This indeed significantly complicates the analysis of equilibrium. For example, it is not trivial to identify which algorithm is replication-proof even under the *single agent setting*.

In what follows, we decompose Theorem 5.3.1 by two separate theorems for single-agent and multi-agent arguments, and provide its proofs as well as the corresponding formal restatement.

Theorem C.4.1 (Single-agent failure - restatement of Theorem 5.3.1)**.** *There exists a problem instance*

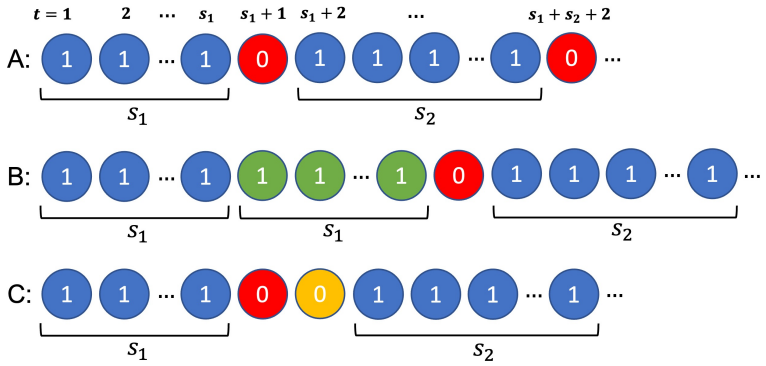


Figure C.1: Bad instance for UCB in Theorem C.4.1. Each node denotes the sequence of the bandit algorithm's selection for each instance. Blue node denotes the original arm with mean reward 1, and green one is the replica of it. Red nodes the original arm with mean reward 0, and yellow node is the replica of it. Note that at round $t = s_1 + s_2 + 2$, instance A realizes reward 0 two times, whereas B and C realizes three times, thus having 1.5 times in average.

such that UCB1 is not replication-proof in the single-agent setting.

Proof of Theorem C.4.1. Consider a single agent with two arms indexed by 1 and 2 from distribution D in which each arm's mean reward is 1 with probability $1/2$ and 0 with probability $1/2$. Consider a truthful strategy $\mathcal{O}$ and strategy $\mathcal{S}$ that replicates arm 1. There exists four possible cases regarding the realization of the arms parameters: (i) both the arms have mean reward 1, (ii) only arm 1 has mean reward 1, (iii) only arm 2 has mean reward 1, and (iv) both the arms have mean reward 0. Denote each cases by X_i for $i = 1, 2, 3, 4$. For X_1 and X_4, note that there exists no difference in the agent's ex-post utility. For X_2 and X_3, we show that there exists a time horizon T such that the average ex-post regret is minimized for strategy $\mathcal{S}$.

Consider the following list of bandit problem instances with Bernoulli arms.

- Let A be a bandit instance with $\mathcal{I}_A = (1, 0)$.
- Let B be a bandit instance with $\mathcal{I}_B = (1, 1', 0)$.
- Let C be a bandit instance with $\mathcal{I}_C = (1, 0, 0')$.

Each element in $\mathcal{I}$ denotes the sequence of mean reward of Bernoulli arms, where we write $1'$ and $0'$ to distinguish the arms with same mean. Importantly, observe that A denotes the ex-post problem instance that the algorithm faces against strategy $\mathcal{O}$ under X_2 or X_3. Furthermore, B denotes the ex-post problem instance against $\mathcal{S}$ under X_2, and C denotes that against $\mathcal{S}$ under X_3.

We now show the following claim.

Claim C.4.2. *Suppose that we run UCB with exploration parameter $f(\cdot)$ for all the problem instances A, B, and C. Then, there exists a time horizon T such that* $\textsc{Reg}_A(T) > {}^{(\textsc{Reg}_B(T)+\textsc{Reg}_C(T))}/_2$.

Proof of the claim. Define $s_1 = 1$. For any integer $i \geq 2$, let s_i denote the minimum integer satisfying

$$1 + \sqrt{\frac{\ln(s_i + i)}{s_i}} < 0 + \sqrt{\frac{\ln(s_i + i)}{i}}.$$

Such s_i exists for any $i \geq 2$ since LHS is a decreasing function of s_i which goes to 1 as s_i increases and RHS is an increasing function of s_i which goes to ∞ as s_i increases, and putting $s_i = i$ yields that LHS is larger than RHS. Further, it is easy to check that $s_i > s_j$ if $i > j$ and $s_i \geq i$.

Due to our construction of the sequence s_i, in A, arm with mean 1 is selected exactly s_i times and then arm with mean 0 is selected for a single round, and it repeats for $i = 1, 2, \ldots$. Figure C.1 illustrates this bandit instances. As can be seen in the figure, the bandit process in A can be described as

$$(1 : s_1, 0 : 1, 1 : s_2, 0 : 1, 1 : s_3, 0 : 1, \ldots),$$

where in each element of $x : y$ denotes that arm with mean x is selected y times. Likewise, in B, we

have

$$(1 : s_1, 1' : s_1, 0 : 1, [1 : s_2, 1' : s_2], 0 : 1, [1 : s_3, 1' : s_3], 0 : 1, \ldots),$$

where for each element in $[\cdot]$, we ignore the orders how 1 and $1'$ are selected, but aggregate them to only present the total number of times each one is selected before getting to the arm 0. For C, we have

$$(1 : s_1, [0 : 1, 0' : 1], 1 : s_2, [0 : 1, 0' : 1], 1 : s_3, [0 : 1, 0' : 1], \ldots).$$

Set $T = s_1 + 2 + s_2$. Given T, in A, the arm 0 is selected twice. In B, the arm 0 is selected once. In C, the arm 0 is selected once, and also $0'$ is selected once. Hence, at round $t = T$, $\textsc{Reg}_A(T) = 2 > \frac{1+2}{2} = \frac{\textsc{Reg}_B(T)+\textsc{Reg}_C(T)}{2}$, and it completes the proof. a

Set T as defined in the claim above. Finally, we can decompose the agent's ex-ante utility as follows.

$$\begin{aligned}
u(\mathcal{O}) = \frac{1}{4}\left(\sum_{i=1}^{4} U(\mathcal{O}; E_i)\right) &= \frac{1}{4}\left(\sum_{i=1,4} U(\mathcal{O}; E_i)\right) + \frac{1}{4}\left(\sum_{i=2,3} U(\mathcal{S}; E_i)\right) \\
&= \frac{1}{4}\left(\sum_{i=1,4}\left(\max_{a\in E_i}\mu_a \cdot T - \textsc{Reg}_{E_i}(T)\right)\right) + \frac{1}{4}\left(\sum_{i=2,3}\left(\max_{a\in E_i}\mu_a \cdot T - \textsc{Reg}_{E_i}(T)\right)\right) \\
&= \frac{T - \textsc{Reg}_A(T)}{2} \\
&< \frac{T - (\textsc{Reg}_B(T) + \textsc{Reg}_C(T))/2}{2} \\
&= \frac{1}{4}\left(\sum_{i=1,4} U(\mathcal{S}; E_i)\right) + \frac{1}{4}\left(\sum_{i=2,3} U(\mathcal{S}; E_i)\right) \\
&= u(\mathcal{S}),
\end{aligned}$$

and thus $\mathcal{S}$ yields the strictly larger ex-ante utility, and it concludes the proof. a

Furthermore, even in the multi-agent setting, we prove that H-UCB cannot be made replication-proof with any choice of algorithm parameters, thereby showing that previously suggested algorithms cannot be made replication-proof for any number of agents.

Theorem C.4.3 (Multi-agent failure - restatement of Theorem 5.3.1)**.** *For any number of agents $n \geq 1$, there exists a problem instance such that H-UCB is not replication-proof.*

Discussion C.4.4. We further note that the above two theorems can be easily generalized by parameterizing the exploration parameters $f(n_i, t) = \sqrt{2\ln(t)/n_i}$ or $g(n_{a,t}, n_i) = \sqrt{2\ln(n_i)/n_{a,t}}$. For example, consider a class of exploration functions f such that it is a monotone decreasing on n_a given t, and a monotone increasing on t given n_a. Further, assume that $f(\cdot, t) \to \infty$ as $t \to \infty$ and $f(n_a, n_a + c) \to 0$ as $n_a \to \infty$ for any constant $c \in \mathbb{N}$, *i.e.,* it decreases more rapidly on n_a than it increases from t. Then, one can check that the both results carry over.

Proof of Theorem C.4.3. If there exists only a single agent, it is easy to see that the dynamics of the algorithm reduces to standard UCB1. Thus, by Theorem C.4.1, H-UCB is not replication-proof. Now we consider multiple agents with $n \geq 2$. We only show for the case $n = 2$ since it easily extends to the cases with $n \geq 3$. Without loss of generality, we assume that the algorithm breaks tie uniformly at random in both phases 1 and 2. Suppose that agent 1 has two deterministic arms sampled from D in which the mean reward is $c \in [0, 1]$ with probability 0.5, and 0 otherwise, for some $c > 0$. We further assume that agent 2 is equipped with a deterministic prior D_i supported on μ, and the arm also deterministically outputs reward μ. Our choice of μ will be determined later, which will be selected in order to propagate the single-agent failure to the two-agent cases. Especially, we will focus on the

round that agent 1 is selected exactly $s_1 + s_2 + 2$ times, as defined in the proof of Theorem C.4.1. Note that the bandit process is deterministic for any realization of the arms' parameters.

We consider a truthful strategy $\mathcal{O}$ of registering both arms, and another strategy $\mathcal{S}$ of replicating only a single arm, and focus on the difference on the agent's expected utility between two strategies. If the agent samples c or 0 both the arms, then the rewards of the bandit process and the agent's ex-post utility therein will be indifferent. Thus, we only focus on the remaining case. Suppose that agent 1 samples c for one arm and 0 for the other, which occurs with probability $1/2$. Similar to the proof of Claim C.4.2, we consnider the following list of bandit problem instances with deterministic arms.

- Let A be a bandit instance with $\mathcal{I}_A = (c, 0)$.
- Let B be a bandit instance with $\mathcal{I}_B = (c, c', 0)$.
- Let C be a bandit instance with $\mathcal{I}_C = (c, 0, 0')$.

Each element in $\mathcal{I}$ denotes the sequence of mean reward of Bernoulli arms, where we write c' and $0'$ to distinguish the arms with same mean. We simply say arm c and arm 0 to denote the arm with the corresponding parameter c and 0, respectively. Then, the realization of the agent 1's arms in playing strategy $\mathcal{S}$ corresponds to either of bandit instance B or C, and strategy $\mathcal{O}$ corresponds to the bandit instance A. Let us abuse the notation of $s_1, s_2, \ldots$ defined in the proof of Claim C.4.2, so that it denotes the number of rounds that arm c is selected consecutively before arm 0 is sampled i-th time, in problem instance A. Also, we use $R_t(i)$ and $N_t(i)$ to denote the quantity of $R(i)$ and $N(i)$ at round t.

Given the random bits of the tie-breaking rule used by the algorithm, let $\tau_1^i, \tau_2^i, \ldots, \tau_k^i, \ldots$ be the rounds that agent 1 is sampled k times when the mean rewards of the agent 1's arms are realized to be $i = A, B, C$. It is straightforward to check that $\tau_t^B \leq \tau_t^A \leq \tau_t^C$[1]. Let $k = s_1 + s_2 + 2$, and following the

[1]This can be shown using the coupling technique introduced in [22].

argument of Claim C.4.2, we observe that at round τ_k^B, agent 1's empirical mean rewards by playing $\mathcal{S}$ will be either of (i) $\frac{c(s_1+s_2)}{s_1+s_2+2}$ or (ii) $\frac{c(s_1+s_2+1)}{s_1+s_2+2}$ based on whether the arms are realized to be instance B or C. By playing $\mathcal{O}$, agent 1 simply has empirical mean reward of $\frac{c(s_1+s_2)}{s_1+s_2+2}$ at round τ_k^A. Furthermore, if $\tau_k^B < \tau_k^A$, then in problem instance A, it means that the agents' parameters at round τ_k^B satisfy

$$R_{\tau_k^B}(1) + \sqrt{\frac{2\ln(t)}{k-1}} < R_{\tau_k^B}(2) + \sqrt{\frac{2\ln(t)}{\tau_k^B - k}},$$

where $R_t(i)$ denotes the empirical average reward of agent i at round t, $R_{\tau_k^B}(1) = \frac{c(s_1+s_2)}{s_1+s_2+2}$ and $R_{\tau_k^B}(2) = \mu$. However, from our construction of the variables and the fact that H-UCB selects agent 1 at round τ_k^B in B, we know that

$$\frac{c(s_1+s_2)}{s_1+s_2+2} + \sqrt{\frac{2\ln(t)}{k-1}} \geq \mu + \sqrt{\frac{2\ln(t)}{\tau_k^B - k}},$$

and thus τ_k^A and τ_k^B needs to be exactly equivalent. Similarly, for instance C, since arm c and c' are selected the same number of times at the beginning of round τ_k^B, we conclude that τ_k^i are equivalent for $i = A, B, C$.

Now define $\theta = \sqrt{\frac{2\ln(t)}{k}} - \sqrt{\frac{2\ln(t)}{\tau_k^B - k}}$ and we pick any $\mu \in (\frac{c(s_1+s_2)}{s_1+s_2+2} + \theta, \frac{c(s_1+s_2+1)}{s_1+s_2+2} + \theta)$.

Then, at round $\tau_k^B + 1$, for bandit instance A and C, we observe that agent 2 will be selected

since

$$R_{\tau_k^B+1}(2) + \sqrt{\frac{2\ln(t)}{N_{\tau_k^B}(2)}} = \mu + \sqrt{\frac{2\ln(t)}{\tau_k^B - k}}$$
$$> \frac{s_1+s_2}{s_1+s_2+2} + \theta + \sqrt{\frac{2\ln(t)}{\tau_k^B - k}}$$
$$= \frac{s_1+s_2}{s_1+s_2+2} + \sqrt{\frac{2\ln(t)}{k}} = R_{\tau_k^B+1}(1) + \sqrt{\frac{2\ln(t)}{N_{\tau_k^B}(1)}}.$$

For bandit instance C, however, observe that

$$R_{\tau_k^C+1}(2) + \sqrt{\frac{2\ln(t)}{N_{\tau_k^C}(2)}} = \mu + \sqrt{\frac{2\ln(t)}{\tau_k^C - k}}$$
$$< \frac{s_1+s_2+1}{s_1+s_2+2} + \theta + \sqrt{\frac{2\ln(t)}{\tau_k^C - k}}$$
$$= \frac{s_1+s_2+1}{s_1+s_2+2} + \sqrt{\frac{2\ln(t)}{k}} = R_{\tau_k^C+1}(1) + \sqrt{\frac{2\ln(t)}{N_{\tau_k^C}(1)}},$$

and thus arm 1 is selected. This essentially implies that by setting $T = \tau_k^A + 1$, agent 1 is selected at round T with probability $1/4$ if he plays $\mathcal{S}$, is never selected if he plays $\mathcal{O}$, where the probability $1/4$ comes from the fact that problem instance C is realized with probability $1/4$. Thus, agent 1's expected utility becomes larger for such T by replicating one of its arm, and we finish the proof.

a

Proof of Theorem 5.4.4. Let $\mathfrak{A}$ be an algorithm with symmetry and TRP. Given a single agent problem instance $\mathcal{I}$ with original set $\mathcal{O}$ with l arms, let $\mathcal{S}$ be a truthful strategy that registers $\mathcal{O}$. Consider any non-truthful strategy $\mathcal{S}$ such that it replicates $r_i \in \mathbb{Z}_0$ times for arm a for $a \in [l]$, where $\mathbb{Z}_0 = \{0, 1, 2, \ldots\}$ denotes the set of nonnegative integers. Note that since $\mathcal{S}$ is non-truthful strategy, at least

one r_a has to be strictly positive for some $a \in [l]$. We abuse $r_a = 0$ to say that arm a is only registered once without any replication, and write $r = (r_1, \ldots, r_l)$. For $v = (v_1, \ldots, v_l) \in \mathbb{R}^l_{\geq 0}$, we write $g(v)$ to denote the density of the mean rewards of arms having v, *i.e.,* $g(v) = \prod_{a=1}^{l} f(v_a)$, where $\mathbb{R}^l_{\geq 0}$ denotes the set of l-dimensional nonnegative real vectors. For simplicity, we assume that the prior distribution of the agent has probability density function, and does not have any point mass.[2] We further write dv to denote $dv_1 dv_2 \ldots dv_l$. We use $\textsc{Reg}_{v,r}$ to denote the regret of problem instance such that the arms' mean rewards are realized to be v and each arm i is replicated r_i times. We finally write $\sigma(v)$ to the permuted vector given a permutation $\sigma \in S^l$, and $\sigma(v)_i$ to denote i-coordinate of $\sigma(v)$.

Observe that the expected utility in playing the strategy $\mathcal{S}$ can be written as follows.

$$\begin{aligned}
u(\mathcal{S}) &= \int_{\mu_1 = v_1} \ldots \int_{\mu_l = v_l} \mathbb{E}\left[U(\mathcal{S}) \mid \bigtimes_{i \in [l]} (v_a : r_a) \right] g(v) dv \\
&= \int_{\mu_1 = v_1} \int_{\mu_2 = v_2 < v_1} \ldots \int_{\mu_k = v_l < v_{l-1}} \sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[U(\mathcal{S}) \mid \bigtimes_{i \in [l]} (v_{\sigma(a)} : r_a) \right] g(v) dv \\
&= \int_{\mu_1 = v_1} \int_{\mu_2 = v_2 < v_1} \ldots \int_{\mu_k = v_l < v_{l-1}} \sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[v_1 T - \textsc{Reg}_{\sigma(v),r}(T) \mid \bigtimes_{i \in [l]} (v_a : r_a) \right] g(v) dv,
\end{aligned}$$

where in the last equation, we condition on $\bigtimes_{i \in [n]} (v_i : r_i)$ since $\textsc{Reg}_{\sigma(v),r}(T)$ already accounts for the permuted bandit instance. Further, due to the permutation invariance of the algorithm, applying σ^{-1} to all the arm indices does not change the expected regret, and thus we can further expand it to be

$$u(\mathcal{S}) = \int_{\mu_1 = v_1} \int_{\mu_2 = v_2 < v_1} \ldots \int_{\mu_l = v_l < v_{l-1}} \sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[v_1 T - \textsc{Reg}_{v,\sigma(r)}(T) \mid \bigtimes_{a \in [l]} (v_a : r_a) \right] g(v) dv. \quad \text{(C.1)}$$

Let $\vec{0}_l$ denote the l-dimensional zero vector. We now prove the following claim.

[2] This assumption is for ease of exposition, and does not restrict any of our results. For instance, if there exists point mass so that two arms may realize the same mean reward with strictly positive probability, then we can essentially unionize those two arms and consider it as a single arm. In this way, one can verify that all the analysis carry over.

Claim C.4.5. *For any $v \in R^l_{\geq 0}$, $\sigma \in \mathcal{P}_l$, and $r \in \mathbb{Z}^l_{\geq 0}$, we have*

$$\sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[\, \textsc{Reg}_{v,\sigma(\vec{0}_l)}(T) \,\right] \leq \sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[\, \textsc{Reg}_{v,\sigma(r)}(T) \,\right].$$

Proof. This directly follows from our definition of TRP. More formally,

$$\begin{aligned}
\sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[\, \textsc{Reg}_{v,\sigma(\vec{0}_l)}(T) \,\right] &= \frac{\text{RP-}\textsc{Reg}(\mathfrak{A}, T; \mathcal{O})}{|\mathcal{P}_l|} && \text{(def. of RP-Regtret)} \\
&\leq \frac{\text{RP-}\textsc{Reg}(\mathfrak{A}, T; \mathcal{S})}{|\mathcal{P}_l|} && \text{(since } \mathfrak{A} \text{ is TRP)} \\
&= \sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[\, \textsc{Reg}_{v,\sigma(r)}(T) \,\right]. && \text{(def. of RP-Regret)}
\end{aligned}$$

It completes the proof. a

By the claim, we can further expand from (C.1) as follows.

$$u(\mathcal{S}) \leq \int_{\mu_1 = v_1} \int_{\mu_2 = v_2 < v_1} \cdots \int_{\mu_l = v_l < v_{l-1}} \sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[v_1 T - \textsc{Reg}_{v,\sigma(\vec{0}_l)}(T) \mid \bigtimes_{a \in [l]} (v_a : r_a) \right] g(v) dv,$$

Since $\textsc{Reg}_{v,\sigma(\vec{0}_l)}$ is equivalent to the regret of the truthful strategy $\mathcal{O}$ given the ex-post realization v, we obtain

$$\begin{aligned}
u(\mathcal{S}) &\leq \int_{\mu_1 = v_1} \int_{\mu_2 = v_2 < v_1} \cdots \int_{\mu_l = v_l < v_{l-1}} \sum_{\sigma \in \mathcal{P}_l} \mathbb{E}\left[v_1 T - \textsc{Reg}_{\mathcal{I}}(T) \mid \bigtimes_{a \in [l]} (v_a : r_a) \right] g(v) dv \\
&= \int_{\mu_1 = v_1} \int_{\mu_2 = v_2} \cdots \int_{\mu_l = v_l} \mathbb{E}\left[U(\mathcal{O}) \mid \bigtimes_{a \in [l]} (v_a : r_a) \right] g(v) dv \\
&= u(\mathcal{O}),
\end{aligned}$$

and thus the truthful registration $\mathcal{O}$ is a dominant strategy. a

C.5 Proof of Proposition 5.4.5 and Theorem 5.4.9

Here, we provide the proofs for the necessity of TRP for truthfulness in the single-agent setting, and the main result in the single-agent setting.

Proof of Proposition 5.4.5. Consider a single-agent setting. Since the algorithm $\mathfrak{A}$ does not satisfy TRP, there exists a problem instance with original arms $\mathcal{O}$ and the agent's strategy $\mathcal{S}$, time horizon T such that

$$\text{RP-REG}(\mathfrak{A}, T; \mathcal{O}) \leq \text{RP-REG}(\mathfrak{A}, T; \mathcal{S}).$$

As observed in the proof of Theorem 5.4.4, the expected utility of the strategies $\mathcal{O}$ and $\mathcal{S}$ can be compared using $\text{RP-REG}(\mathfrak{A}, T; \mathcal{O})$ and $\text{RP-REG}(\mathfrak{A}, T; \mathcal{S})$, so the result follows. a

Before presenting the proof of Theorem 5.4.9, we provide the pseudocode of the standard ETC algorithm to make the paper self-contained.

ALGORITHM 7: ETC
Input: exploration length m, tie-breaking rule
Play each arm for m rounds.
Play the arm with the highest empirical mean for all remaining rounds

Proof of Theorem 5.4.9. Recall that we will deal with a single agent. Consider a bandit instance $\mathcal{I}$ in which there exists l arms with parameters $v_1 \geq v_2 \geq \ldots \geq v_l$, and the time horizon is T. Let $n_a(t)$ be the number of rounds that arm a is pulled and $\hat{\mu}_a(t)$ be the empirical average reward, at the end of round t, and define $\delta_a = v_1 - v_a$ for $a \in [l]$.

Our proof mainly relies on the following lemma.

Lemma C.5.1. *Suppose that the realization of the arms's mean rewards is given, and let a^* be the arm the largest mean reward. If $m \geq \frac{2l}{\Delta^2} \ln 2T$, then for a sub-optimal arm we have:*

$$\mathbb{E}\left[\, n_a(T) \,\right] \leq m + 1.$$

Especially for $m = \frac{2}{\Delta^2} l \ln(2T)$, the expected regret is upper bounded by:

$$\textsc{Reg}(T) \leq \sum_{a \in [l]} (m+1)\Delta_i \leq \sum_{a \in [l]} \Big(\frac{2\delta_a l \ln(2T)}{\Delta^2} + 1 \Big).$$

Proof of the lemma. Due to the structural nature of ETC, for any sub-optimal arm a we have

$$\mathbb{E}\left[\, n_a(T) \,\right] \leq m + (T - mk) \Pr\left[\, \hat{\mu}_a(m) \geq \hat{\mu}_{a^*}(m) \,\right]$$

Now we focus on upper bounding $\Pr\left[\, \hat{\mu}_a(m) \geq \hat{\mu}_{a^*}(m) \,\right]$.

Claim C.5.2. *For any arm $a \neq a^*$, the following holds.*

$$\Pr\left[\, \hat{\mu}_a(m) \geq \hat{\mu}_{a^*}(m) \,\right] \leq 2 \exp(-\frac{\delta_a^2 m}{2})$$

Proof of the claim. We can expand the inequality as follows.

$$\begin{aligned}
\Pr\left[\,\hat{\mu}_a(m) \geq \hat{\mu}_{a^*}(m)\,\right] &= \Pr\left[\,(\hat{\mu}_a(m) - v_a) + (v_{a^*} - \hat{\mu}_{a^*}(m)) \geq \delta_a\,\right] \\
&\leq \Pr\left[\,(\hat{\mu}_a(m) - v_a) \geq \frac{\delta_a}{2}\,\right] + \Pr\left[\,(v_{a^*} - \hat{\mu}_{a^*}(m)) \geq \frac{\delta_a}{2}\,\right] \\
&\leq 2\exp(-\frac{\delta_a^2}{4}\frac{1}{2}\frac{1}{\frac{1}{4m}}) && \text{(Hoeffding's inequality)} \\
&\leq 2\exp(-\frac{\delta_a^2 m}{2}),
\end{aligned}$$

which finishes the proof. Note that we use the following well-known type of Hoeffding's inequality.

Theorem C.5.3. *Let $X_1, \ldots, X_n$ be independent random variables such that $a_i \leq X_i \leq b_i$ almost surely. Define $S_n = X_1 + \ldots X_n$. For any $t > 0$,*

$$\Pr\left[\,|S_n - \mathbb{E}\left[\,S_n\,\right]| \geq t\,\right] \leq 2\exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

a

Using the claim above, we can further expand,

$$\begin{aligned}
\mathbb{E}\left[\,n_a(T)\,\right] &\leq m + (T - mk)\Pr\left[\,\hat{\mu}_a(m) \geq \hat{\mu}_{a^*}(m)\,\right] \\
&\leq m + 2T\exp(-\frac{\delta_a^2 m}{2}) && \text{(Claim C.5.2 \& } T - mk \leq T) \\
&\leq m + 2T\exp(-\left(\frac{\delta_a}{\Delta}\right)^2 \ln 2T) && (m \geq 2/\Delta^2 \ln(2T)) \\
&\leq m + 2T\exp(-\ln 2T) && \text{(Since } \tfrac{\delta_a}{\Delta} \geq 1) \\
&= m + 1.
\end{aligned}$$

This proves the upper bound on $\mathbb{E}[\, n_a(T)\,]$. To obtain the upper bound on the regret, one can use the following well-known regret decomposition lemma.

Lemma C.5.4 (Regret decomposition)**.** *The regret can be writtten as* $\textsc{Reg}(T) = \sum_{a\in[l]} \mathbb{E}[\, n_i(T)\delta_a\,]$.

From this, the regret bound easily follows as below

$$\begin{aligned}
\textsc{Reg}(T) &\leq \sum_{a\in[l]} (m+1)\delta_a \\
&= \sum_{a\in[l]} (\frac{2l}{\Delta^2}\ln 2T + 1)\delta_a \\
&\leq \sum_{a\in[l]} (\frac{2l\delta_a}{\Delta^2}\ln 2T + 1). && \text{(Since } \Delta_a \leq 1\text{)}
\end{aligned}$$

This completes the proof. a

Now, consider a strategy $\mathcal{S} \neq \mathcal{O}$ of which the replication vector is $r = (r_1, \ldots, r_l)$ (as defined in the proof of Theorem 5.4.4), and at least one coordinate being strictly positive. Consider a realization $v \in \mathbb{R}^l_{\geq 0}$, and a permutation $\sigma(v)$. Let $\mathcal{I}_\sigma$ be a bandit instance in which the arms are permuted by σ from $\mathcal{I}$.

Given $v_1 \geq \ldots \geq v_l$, let $\mathcal{I}_{\sigma,v}$ be the instance in which i-th arm has mean reward $v_{\sigma(a)}$ for $a \in [l]$, *i.e.*, $\mathcal{I}_{\sigma,v} = \bigtimes_{a\in[l]}(v_{\sigma(a)} : r_a)$. Denote the expected regret under $\mathcal{I}_{\sigma,v}$ by $\textsc{Reg}_{I_{\sigma,v}}$.

From the regret decomposition lemma, we have

$$\textsc{Reg}_{I_{\sigma,v}}(T) = \sum_{a\in[l]} \mathbb{E}\left[\, n_{\sigma(a),t}\,\right]\delta_{\sigma(a)}(r_a+1).$$

Due to the initial exploration phase of ETC, obviously $\mathbb{E}\left[\, n_{\sigma(a),t} \,\right] \geq m$. Hence we obtain

$$\textsc{Reg}_{I_{\sigma,v}}(T) \geq \sum_{a\in[l]} m \cdot \delta_{\sigma(a)}(r_a + 1).$$

Now, RHS in the random permutation regret (5.1) can be written as follows.

$$\begin{aligned}
\sum_{\sigma\in\mathcal{P}_l} \textsc{Reg}_{\mathcal{I}_{\sigma,v}}(T)/|\mathcal{P}_l| &\geq \frac{\sum_{\sigma\in\mathcal{P}_l}\sum_{a\in[l]} m\cdot\delta_{\sigma(a)}(r_a+1)}{l!} \\
&\geq \frac{\sum_{a\in[l]}\sum_{\sigma\in\mathcal{P}_l} m\cdot\delta_{\sigma(a)}(r_a+1)}{l!} && \text{(Change order of summation)} \\
&= \frac{\sum_{a\in[l]}\sum_{\sigma\in\mathcal{P}_l} m\cdot\delta_{a}(r_{\sigma(a)}+1)}{l!} && \text{(Sum over } \sigma \text{ equals that over } \sigma^{-1}\text{)} \\
&= \frac{\sum_{a\in[l]} m\cdot\delta_a (l-1)!\sum_{j=1}^{l}(r_j+1)}{l!} && \text{(fix } \sigma(a)=j \text{ and sum over } \sigma\in\mathcal{P}_l\text{)} \\
&= \frac{\sum_{a\in[l]} m\cdot\delta_a (\|r\|_1+l)}{l} && \text{(C.2)}
\end{aligned}$$

Due to our construction of r, we have $\|r\|_1 \geq 1$, and finally we obtain

$$\begin{aligned}
\textsc{Reg}_{\mathcal{I}}(T) &\leq \sum_{a\in[l]}(m+1)\delta_a && \text{(C.3)} \\
&\leq \sum_{a\in[l]} m\delta_a \frac{l+1}{l}, && (m = \tfrac{2l}{\Delta^2}\ln 2T \geq l) \\
&\leq \frac{\sum_{a\in[l]} m\cdot\delta_a\sum_{j=1}^{l}(r_j+1)}{l} && (\|r\|_1 \geq 1) \\
&\leq \sum_{\sigma\in\mathcal{P}_l} \textsc{Reg}_{\mathcal{I}_{\sigma,v}}(T)/|\mathcal{P}_l|,
\end{aligned}$$

where the last inequality follows from the inequality (C.2). Note further that $\textsc{Reg}_{\mathcal{I}}(T)$ is RP-$\textsc{Reg}$ of truthful strategy $\mathcal{O}$ given the realization v, and RHS denotes that of a strategy $\mathcal{S}$ whose replication

vector is r. Thus,

$$\text{RP-REG}(\mathfrak{A}, T, \mathcal{O}) \leq \text{RP-REG}(\mathfrak{A}, T, \mathcal{S}),$$

for $\mathfrak{A}$ being ETC, and it concludes that ETC is TRP. a

C.6 Proof Sketch of Proposition 5.4.10

We here provide the proof sketch of Proposition 5.4.10, where we omit the details as it is a cumbersome application of standard techniques. The formal pseudocode is provided in Algorithm 8.

ALGORITHM 8: Prior-indepdenent ETC

Input: Tie-break rule τ
$T_{prev} \leftarrow 0$
for $i = 1, 2, \ldots$ **do**
  $T \leftarrow 2^i$
  while $t \leq T$ **do**
    $m \leftarrow \lfloor \ln^4(T) \rfloor$ Play each arm m rounds
    Play the arm with the highest empirical mean for all remaining rounds, given τ
  $T_{prev} \leftarrow T$

To prove the proposition, let us fix some T and focus on the rounds that operates based on this T. Due to the assumptions, there exist sufficiently large T such that $l, 1/\Delta \leq \ln T$. Then, for sufficiently large T, it follows that

$$m = \lfloor \ln^4(T) \rfloor \geq l \ln T \Delta^2. \tag{C.4}$$

If T is not sufficiently large, the while loop will break before gets to the exploitation phase, and therefore it does not change the expected payoff of the agent. On the other hand, if T is sufficiently large, then since m satisfies (C.4), by Theorem 5.4.9, each partition will be replication-proof for fixed

T. Therefore, iterating over $T = 2^1, 2^2, \ldots$ preserves the replication-proofness of the algorithm, and thus the overall algorithm is replication-proof. Note further that the exploration length $\ln^4(T)$ only amounts for polylogarithmic factor to the regret of the algorithm.

ALGORITHM 9: Hierarchical ETC with Restarting (H-ETC-R)

Input: Tie-breaking rule, agent set $\mathcal{N}$, arm set $\mathcal{S}_i$ for $i \in \mathcal{N}$, restarting round $\tau \in [T]$
Let $\hat{\mu}_i \leftarrow 0$, $n_i \leftarrow 0$, $\hat{\mu}_{i,a} \leftarrow 0$, $n_{i,a} \leftarrow 0$ for $i \in \mathcal{N}$ and $a \in \mathcal{S}_i$
for $t = 1, 2, \ldots, T$ **do**
  if $t = \tau + 1$ **then**
    $\hat{\mu}_{i,a} \leftarrow 0$; $n_{i,a} \leftarrow 0$ for $\forall a \in \mathcal{S}_i, \forall i \in \mathcal{N}$
  if $n_i < M$ *for some* $i \in \mathcal{N}$ **then**
    $\hat{i} \leftarrow i$
  else
    $\hat{i} \leftarrow \operatorname{argmax}_{i \in \mathcal{N}} \hat{\mu}_i$
  if $n_{\hat{i},a} < m$ *for some* $a \in \mathcal{S}_{\hat{i}}$ **then**
    $\hat{a} \leftarrow a$
  else
    $\hat{a} \leftarrow \operatorname{argmax}_{a \in \mathcal{S}_{\hat{i}}} \hat{\mu}_{\hat{i},a}$
  Pull $\hat{a}$ of agent $\hat{i}$ and obtain reward R_t
  Update statistics: $\hat{\mu}_{\hat{i},\hat{a}} \leftarrow \frac{\hat{\mu}_{\hat{i},\hat{a}} n_{\hat{i},\hat{a}} + R_t}{n_{\hat{i},\hat{a}} + 1}$; $\hat{\mu}_{\hat{i}} \leftarrow \frac{\hat{\mu}_{\hat{i}} n_{\hat{i}} + R_t}{n_{\hat{i}} + 1}$; $n_{\hat{i},\hat{a}} \leftarrow n_{\hat{i},\hat{a}} + 1$; $n_{\hat{i}} \leftarrow n_{\hat{i}} + 1$;

C.7 Proof of Theorem 5.5.1

Proof. First, note that our H-ETC-R is well-defined since our choice of M and m ensures that the exploitation phase of the first phase starts after the exploitation of the second phase ends. For simplicity, denote the first phase ETC by ALG_1 and the second phase ETC by ALG_2. Pick an agent i with arms $\{o_{i,1}, \ldots, o_{i,l}\}$, let $\mathcal{S}_i$ be agent i's arbitrary strategy which replicates arm $o_{i,j}$ of $r_{i,j} \in \mathbb{N}_0$ times for each $j \in [l_i]$. Let $r_i = (r_{i,1}, \ldots, r_{i,l_i})$, and denote by $v_i = (v_{i,1}, \ldots, v_{i,l_i})$ the realization of agent i's original arms' mean rewards. Fix strategies of all the other agents $j \neq i$ by $\mathcal{S}_{-i}$, and fix a realization v_{-i} of all the other agents' arms' mean rewards. Henceforth, we will omit $\mathcal{S}_{-i}$ and v_{-i} from the arguments in the notations presented hereafter, unless confusion arises. In addition, we assume that all the prior

distributions do not have point mass for simplicity, but it can easily be generalized as discussed in the proof of Theorem 5.4.4.

Let i^* denote the agent selected at the exploitation phase of ALG_1, and l_i to denote agent i's number of arms. We further write $\hat{\mu}_i(M)$ to denote the empirical average reward of agent i when agent i is selected M times, and $\hat{\mu}_{i,a}(m)$ to denote the empirical average reward of agent i's arm a when it's selected m times. Then, given v_i and r_i, recall that the bandit instance agent i submits can be presented as $\bigtimes_{i\in[l]}(v_a : r_a)$.

Now, similar to the proof of Theorem 5.4.4, we first write agent i's expected utility as follows.

$$u_i(\mathcal{S}_i) = \int_{\mu_{i,1}=v_{i,1}} \int_{\mu_{i,2}=v_{i,2}<v_{i,1}} \dots \int_{\mu_{i,l_i}=v_{i,l_i}<v_{i,l_i-1}} \sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[U_i(\mathcal{S}_i) \mid \bigtimes_{a\in[l_i]}(v_{i,\sigma(a)} : r_{i,a})\right] g(v_i)dv_i. \tag{C.5}$$

We first characterize the expected utility of agent i based on its utility under the single agent setting with agent i. To this end, we introduce several notations. Write $\Gamma_i^{\mathcal{S}_i}(t, \sigma(v_i))$ to denote agent i's utility when he is selected t times under the single agent setting in ALG_2 when its arms' mean rewards are realized to be $\sigma(v_i)$, and he plays strategy $\mathcal{S}_i$ with replication vector r_i. Given M, m, and $\mathcal{S}_i$, define an event

$$C_i^{\mathcal{S}_i}(y, \sigma(v)) = \{\max_{\mathcal{N}\ni j\neq i}(\hat{\mu}_j(M)) = y \mid \bigtimes_{a\in[l_i]}(v_{i,\sigma(a)} : r_{i,a})\},$$

i.e., the event that agents other than i have maximal agent-wise empirical average reward y at the end of

the exploration phase in ALG_1. Again given M, m, and $\mathcal{S}_i$, we further define following two quantities:

$$p_i^{\mathcal{S}_i}(y, \sigma(v_i)) = \Pr\left[C_i(y) \mid \bigtimes_{a \in [l_i]} (v_{i,\sigma(a)} : r_{i,a}) \right]$$
$$q_i^{\mathcal{S}_i}(y, \sigma(v_i)) = \Pr\left[\hat{\mu}_i(M) \geq y \mid \bigtimes_{a \in [l_i]} (v_{i,\sigma(a)} : r_{i,a}) \right].$$

Remark that all the quantities above are conditioned on the event that v_{-i} is realized and the other agents are playing $\mathcal{S}_{-i}$.

The following claim essentially characterizes the expected utility of agent i under strategy $\mathcal{S}$.

Claim C.7.1. *Agent i's expected utility can be written as follows.*

$$\mathbb{E}\left[U_i(\mathcal{S}_i) \mid \bigtimes_{a \in [l_i]} (v_{i,\sigma(a)} : r_{i,a} \right] = \mathbb{E}\left[\Gamma_i^{\mathcal{S}_i}(M, \sigma(v_i)) \right] + \mathbb{E}\left[\Gamma_i^{\mathcal{S}_i}(T - Mn, \sigma(v_i)) \right] \int_0^1 p_i^{\mathcal{S}_i}(y, \sigma(v_i)) q_i^{\mathcal{S}_i}(y, \sigma(v_i)) dy. \tag{C.6}$$

Proof of the claim. Since H-ETC-R is running with the first phase being ETC, the expected utility of agent i can be decomposed as (i) expected utility in the exploration phase and (ii) expected utility in the exploitation phase. Note that (i) essentially coincides with the expectation of $\Gamma_i^{\mathcal{S}_i}(M, \sigma(v_i))$ since it chooses agent i exactly M times regardless of the random rewards. Moreover, due to our restarting at time $\tau = Mn$, when the exploitation phase of ALG_1 starts, all the parameters of ALG_2 is initialized. Thus, conditioned on the event that agent i has the largest empirical average reward at round $\tau = Mn$, its expected utility in the exploitation phase again coincides with the single agent case expected utility $\Gamma_i^{\mathcal{S}_i}(T - Mn, \sigma(v_i))$. Otherwise, if agent i has not been selected as the empirically best agent, then its conditional expected utility is simply zero. Note that the probability of agent i being the empirically

best agent can be written as follows,[3]

$$\begin{aligned}\Pr\left[\,\text{agent } i \text{ is the empirically best agent at round } \tau = Mn\,\right] &= \int_0^1 \Pr\left[\,\hat{\mu}_i(M) \geq y\,\right]\Pr\left[\,C_i(y)\,\right] \\ &= \int_0^1 p_i^{\mathcal{S}_i}(y, \sigma(v_i))q_i^{\mathcal{S}_i}(y, \sigma(v_i))dy,\end{aligned}$$

and we finish the proof of the claim. a

Using the claim above, for any strategy $\mathcal{S}_i$, we can further expand its conditional expected utility $\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\, u_i(\mathcal{S}_i) \mid \bigtimes_{a\in[l_i]}(v_{i,\sigma(a)} : r_{i,a})\,\right]$ as follows.

$$\begin{aligned}&\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\, U_i(\mathcal{S}_i) \mid \bigtimes_{a\in[l_i]}(v_{i,\sigma(a)} : r_{i,a})\,\right] \\ &\quad= \sum_{\sigma\in\mathcal{P}_{l_i}} \left(\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}_i}(M, \sigma(v_i))\,\right] + \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}_i}(T - Mn, \sigma(v_i))\,\right]\int_0^1 p_i^{\mathcal{S}_i}(y, \sigma(v))q_i^{\mathcal{S}_i}(y, \sigma(v_i))dy\right) \\ &\quad= \left(\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(M, \sigma(v_i))\,\right]\right) + \sum_{\sigma\in\mathcal{P}_{l_i}} \left(\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}_i}(T - Mn, \sigma(v_i))\,\right]\int_0^1 p_i^{\mathcal{S}_i}(y, \sigma(v_i))q_i^{\mathcal{S}_i}(y, \sigma(v_i))dy\right) \\ &\quad= \left(\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}_i}(M, \sigma(v_i))\,\right]\right) + \int_0^1 \left(\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}_i}(T - Mn, \sigma(v_i))\,\right]p_i^{\mathcal{S}_i}(y, \sigma(v_i))q_i^{\mathcal{S}_i}(y, \sigma(v_i))\right)dy \\ &\quad= \left(\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}_i}(M, \sigma(v_i))\,\right]\right) + \int_0^1 p_i^{\mathcal{S}_i}(y, \sigma(v_i))\left(\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}_i}(T - Mn, \sigma(v_i))\,\right]q_i^{\mathcal{S}_i}(y, \sigma(v_i))\right)dy,\end{aligned} \tag{C.7}$$

where the third equality follows from Fubini's theorem, and the last inequality holds since $p_i^{\mathcal{S}_i}(y, \sigma(v_i))$ is independent from σ.

In case of the truthful strategy $\mathcal{O}_i$, since any permutation σ yields the same problem instance, all

[3]To be precise, we need to take account for the tie-breaking rule and the event that tie happens, but we omit the detail as it can be easily handled and is beyond of our interest.

the terms inside (C.7) remains the same. Thus,

$$\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[U_i(\mathcal{O}_i) \mid \bigtimes_{a\in[l_i]} (v_{i,\sigma(a)} : r_{i,a}) \right]$$
$$= |\mathcal{P}_{l_i}| \left(\mathbb{E}\left[\Gamma_i^{\mathcal{O}_i}(M, v_i) \right] + \mathbb{E}\left[\Gamma_i^{\mathcal{O}_i}(T - Mn, v_i) \right] \int_0^1 p_i^{\mathcal{O}_i}(y, v_i) q_i^{\mathcal{O}_i}(y, v_i) dy \right). \tag{C.8}$$

Now it suffices to prove that the RHS of (C.8) is at least that of (C.7).

We first observe that the first term can effectively compared based on our result on the single-agent setting.

Claim C.7.2. *For any strategy $\mathcal{S}_i$, we have*

$$\sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\Gamma_i^{\mathcal{O}_i}(M, \sigma(v_i)) \right] \geq \sum_{\sigma\in\mathcal{P}_{l_i}} \mathbb{E}\left[\Gamma_i^{\mathcal{S}_i}(M, \sigma(v_i)) \right].$$

Proof. Note that $\Gamma_i^{\mathcal{O}_i}(M, \sigma(v_i))$ denotes the agent i's utility at round M under the single agent setting. Thus, the claim directly follows from Theorem 5.4.9 since we run ETC with $m = 2l/\Delta^2 \ln(2T)$ as Phase 2 algorithm. a

Thus, it remains to compare the latter term that involves the integral. Given $v_i \in \mathbb{R}_{\geq 0}^{l_i}$, for $a \in [l_i]$, write $Y_{i,a}$ to denote the event that the arm with mean v_a of agent i is selected in the exploitation rounds of ALG_2 for agent i, *i.e.,* having the largest empirical average reward at the end of the exploration rounds in ALG_2 for agent i.

Further, similar to the definition of $\Gamma_i^{\mathcal{S}_i}(\cdot,\cdot)$, we analogously define $\Gamma_{i,a}^{\mathcal{S}_i}(\cdot,\cdot)$ to be the random cumulative rewards when there exists only agent i and only arm a under ALG_2. That is, $\Gamma_{i,a}^{\mathcal{S}_i}(t, \sigma(v_i))$ is just a summation of t random variables where each random variable follows the agent i's arm's reward

distribution parameterized by the mean reward $\sigma(v_i)$. Note that our construction in equation (C.5) verifies $v_{i,1} \geq v_{i,2} \geq \ldots \geq v_{i,l_i}$. Define $\delta_{i,a} = v_{i,1} - v_{i,a}$ for $a \in [l_i]$.

We now construct a lower bound of $q_i^{\mathcal{O}_i}(y, v_i)$ in the following.

Claim C.7.3. *For any $y \in [0,1]$, and $\sigma \in \mathcal{P}_{l_i}$, consider arbitrary $v_i \in [0,1]^{l_i}$ such that $v_{i,1} \geq v_{i,2} \geq \ldots \geq v_{i,l_i}$. Then, we have*

$$q_i^{\mathcal{O}_i}(y, v_i) \geq \Pr\left[\, \Gamma_{i,1}^{\mathcal{O}_i}(M, v_i) \geq My \,\right] \left(1 - 2\sum_{a=2}^{l_i} \exp(-\frac{\delta_{i,a}^2 m}{2})\right)$$

Proof of the claim. Due to the independence of rewards from arm a at exploitation phase and the random variable $\mathbf{1}\left\{ Y_{i,a} \right\}$, we can rewrite $q_i^{\mathcal{O}}(y, \sigma(v))$ as follows.

$$\begin{aligned}
q_i^{\mathcal{O}_i}(y, v_i) &= \Pr\left[\sum_{a=1}^{l_i} \left(\Gamma_{i,a}^{\mathcal{O}_i}(m, v) + \mathbf{1}\left\{ Y_{i,a} \right\} \Gamma_{i,a}^{\mathcal{O}_i}(M - l_i m, v_i) \right) \geq My \right] \\
&= \sum_{a=1}^{l_i} \Pr\left[\sum_{a=1}^{l_i} \left(\Gamma_{i,a}^{\mathcal{O}}(m, v) + \mathbf{1}\left\{ Y_{i,a} \right\} \Gamma_{i,a}^{\mathcal{O}}(M - l_i m, v) \right) \geq My \mid Y_{i,a} \right] \Pr\left[Y_{i,a} \right] \\
&= \sum_{a=1}^{l_i} \Pr\left[\left(\sum_{a=1}^{l_i} \Gamma_{i,a}^{\mathcal{O}}(m, v) \right) + \Gamma_{i,a}^{\mathcal{O}}(M - l_i m, v) \geq My \mid Y_{i,a} \right] \Pr\left[Y_{i,a} \right] \\
&\geq \Pr\left[\left(\sum_{a=1}^{l_i} \Gamma_{i,1}^{\mathcal{O}}(m, v) \right) + \Gamma_{i,1}^{\mathcal{O}}(M - l_i m, v) \geq My \mid Y_{i,1} \right] \Pr\left[Y_{i,1} \right]
\end{aligned}$$

Note further that due to the positive correlation between the event $Y_{i,1}$ and the tail event, it follows that

$$\begin{aligned}
&\Pr\left[\left(\sum_{a=1}^{l_i} \Gamma_{i,1}^{\mathcal{O}_i}(m, v) \right) + \Gamma_{i,1}^{\mathcal{O}_i}(M - l_i m, v) \geq My \mid Y_{i,1} \right] \\
&\geq \Pr\left[\left(\sum_{a=1}^{l_i} \Gamma_{i,1}^{\mathcal{O}_i}(m, v) \right) + \Gamma_{i,1}^{\mathcal{O}_i}(M - l_i m, v) \geq My \right].
\end{aligned}$$

Let us write

$$\psi_i(M, y, v) = \Pr\left[\left(\sum_{a=1}^{l_i} \Gamma_{i,1}^{\mathcal{O}_i}(m, v)\right) + \Gamma_{i,1}^{\mathcal{O}_i}(M - l_i m, v) \geq My\right],$$

then we have

$$q_i^{\mathcal{O}_i}(y, v_i) \geq \psi_i(M, y, v)\Pr\left[Y_{i,1}\right] \tag{C.9}$$

Since $Z_{i,1}$ is the complement of $\cup_{a=2}^{l} Z_{i,a}$, observe that

$$\begin{aligned}
\Pr\left[Z_{i,1}^c\right] &= \Pr\left[\cup_{a=2}^{l} Z_{i,a}\right] \\
&\leq \sum_{a=2}^{l} \Pr\left[Z_{i,a}\right] && \text{(Union bound)} \\
&\leq \sum_{a=2}^{l} \Pr\left[\hat{\mu}_{i,1}(m) \leq \hat{\mu}_{1,a}(m)\right] && \text{(Definition of } Z_{i,a}\text{)} \\
&\leq \sum_{a=2}^{l} 2\exp(-\frac{\delta_{i,a}^2 m}{2}). && \text{(Claim C.5.2)}
\end{aligned}$$

Hence, we obtain

$$q_i^{\mathcal{O}}(y, \sigma(v)) \geq \psi_i(M, y, v)\left(1 - 2\sum_{a=2}^{l} \exp(-\frac{\delta_{i,a}^2 m}{2})\right),$$

and it proves the claim. a

Claim C.7.4. *For any strategy $\mathcal{S}_i$, $\sigma \in S_{l_i}$, $y \in [0, 1]$ and $v \in [0, 1]^{l_i}$, we have*

$$q_i^{\mathcal{S}_i}(y, \sigma(v)) \leq \psi_i(M, y, v)$$

Proof of the claim. This easily follows from the fact that in the probability $\psi(M, y, v)$, the exploration phase runs with the original l_i arms for m times each and then the exploration phase runs with the optimal arm $(i, 1)$, whereas in the probability $q_i^{\mathcal{S}_i}(y, \sigma(v))$, the exploration runs with the original l_i arms upon the replicated arm for m times each, and then exploitation runs with the empirically best arm. That is, we can couple the first $l_i \cdot m$ rounds for the exploration phase of the original l_i arms, and then the exploitation is dominant for $\psi(M, y, v)$ due to the stochastic dominance of the arm $(i, 1)$ over the other arms.

a

Now, combining Claim C.7.3 and C.7.4, we obtain

$$\begin{aligned}
q_i^{\mathcal{O}}(y, \sigma(v)) - q_i^{\mathcal{S}}(y, \sigma(v)) &\geq -2\psi_i(M, y, v)\left(\sum_{a=2}^{l_i} \exp(-\frac{\delta_{i,a}^2 m}{2})\right) \\
&\geq -2l_i\psi_i(M, y, v)\left(\exp(-\frac{\Delta^2 m}{2})\right) && \text{(Due to } \delta_{i,a} \geq \Delta\text{)} \\
&\geq -2l_i \cdot \frac{q_i^{\mathcal{O}}(y, v)}{\Pr\left[\, Y_{i,1} \,\right]} \cdot \left(\exp(-\frac{\Delta^2 m}{2})\right) && \text{(Due to ineq. (C.9))} \\
&\geq -2l_i \cdot \frac{q_i^{\mathcal{O}}(y, v)}{1 - \sum_{a=2}^{l_i} 2\exp(-\frac{\delta_i^2 m}{2})} \cdot \left(\exp(-\frac{\Delta^2 m}{2})\right) \\
&\geq -2l_i \cdot \frac{q_i^{\mathcal{O}}(y, v)}{1 - 2l_i \exp(-\frac{\Delta^2 m}{2})} \cdot \left(\exp(-\frac{\Delta^2 m}{2})\right), && \text{(C.10)}
\end{aligned}$$

where the fourth inequality follows from the concentration on $Z_{i,1}$ in the proof of Claim C.7.3.

Now we consider two cases regarding σ: (i) when $\sigma(1) = 1$ and (ii) otherwise.

Recall that it suffices to prove that the RHS of (C.8) is at least that of (C.7). Due to Claim C.7.2,

it suffices to prove the following inequality.

$$\int_0^1 p_i^{\mathcal{S}}(y,\sigma(v))\left(\sum_{\sigma\in S_l}\underbrace{\left\{\mathbb{E}\left[\,\Gamma_i^{\mathcal{O}}(T-Mn,\sigma(v))\,\right]q_i^{\mathcal{O}}(y,\sigma(v))-\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]q_i^{\mathcal{S}}(y,\sigma(v))\right\}}_{\Theta(\sigma,v)}\right)dy\geq 0. \tag{C.11}$$

Note that the term inside the summation can be decomposed as follows.

$$\begin{aligned}\Theta(\sigma,v)=&\left\{\mathbb{E}\left[\,\Gamma_i^{\mathcal{O}}(T-Mn,\sigma(v))\,\right]-\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\right\}q_i^{\mathcal{O}}(y,\sigma(v))\\&+\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\left\{q_i^{\mathcal{O}}(y,\sigma(v))-q_i^{\mathcal{S}}(y,\sigma(v))\right\}.\end{aligned}$$

Due to (C.10), we further obtain

$$\begin{aligned}\Theta(\sigma,v)\geq&\underbrace{\left\{\mathbb{E}\left[\,\Gamma_i^{\mathcal{O}}(T-Mn,\sigma(v))\,\right]-\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\right\}q_i^{\mathcal{O}}(y,\sigma(v))}_{\Theta_1(\sigma,v)}\\&\underbrace{-2\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\cdot l_i\cdot\frac{q_i^{\mathcal{O}}(y,v)}{1-2l_i\exp(-\frac{\Delta^2 m}{2})}\cdot\left(\exp(-\frac{\Delta^2 m}{2})\right)}_{\Theta_2(\sigma,v)}.\end{aligned}$$

By summing over $\sigma\in S_{l_i}$, for $\Theta_1(\sigma,v)$, we have

$$\begin{aligned}\sum_{\sigma\in S_l}\Theta_1(\sigma,v)&=\sum_{\sigma\in S_l}\left\{\mathbb{E}\left[\,\Gamma_i^{\mathcal{O}}(T-Mn,\sigma(v))\,\right]-\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\right\}q_i^{\mathcal{O}}(y,\sigma(v))\\&=q_i^{\mathcal{O}}(y,\sigma(v))\cdot\sum_{\sigma\in S_{l_i}}\left\{\mathbb{E}\left[\,\Gamma_i^{\mathcal{O}}(T-Mn,\sigma(v))\,\right]-\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\right\}\end{aligned}$$

From inequality (C.2) and (C.3) we know that

$$\begin{aligned}\sum_{\sigma\in S_{l_i}}\left\{\mathbb{E}\left[\,\Gamma_i^{\mathcal{O}}(T-Mn,\sigma(v))\,\right]-\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\right\} &\geq \sum_{a\in[l_i]}\delta_{i,a}(\frac{m}{l_i}-1)\\ &\geq \sum_{a\in[l_i]}\delta_{i,a}(\frac{2l_i\ln 2T/\Delta^2}{l_i}-1)\\ &= \sum_{a\in[l_i]}\delta_{i,a}(\frac{\ln 2T}{\Delta^2}-1),\end{aligned}$$

which implies

$$\sum_{\sigma\in S_l}\Theta_1(\sigma,v)\geq q_i^{\mathcal{O}}(y,\sigma(v))\cdot\sum_{i\in[l]}\delta_{i,a}(\frac{\ln 2T}{\Delta^2}-1).$$

Putting $m=4\cdot\frac{2}{\Delta^2}(\max_i l_i)\ln(2T)$, observe that the following holds.

$$\exp(-\frac{\Delta^2 m}{2})\leq\frac{1}{4l_iT}$$

Hence, for $\Theta_2(\sigma,v)$, we further obtain

$$\begin{aligned}\sum_{\sigma\in S_{l_i}}\Theta_2(\sigma,v) &= 2\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\cdot l_i\cdot\frac{q_i^{\mathcal{O}}(y,v)}{1-2l_i\exp(-\frac{\Delta^2 m}{2})}\cdot\left(\exp(-\frac{\Delta^2 m}{2})\right)\\ &\leq 4\mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\cdot l_i\cdot q_i^{\mathcal{O}}(y,v)\cdot\frac{1}{4l_iT}\\ &= \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]q_i^{\mathcal{O}}(y,v)\frac{1}{4T}\\ &\leq \frac{q_i^{\mathcal{O}}(y,v)}{4} \qquad \text{(Since } \mathbb{E}\left[\,\Gamma_i^{\mathcal{S}}(T-Mn,\sigma(v))\,\right]\leq T\text{)}\end{aligned}$$

Combining the above observations, for $\Theta(\sigma, v)$, we obtain

$$\begin{aligned}\sum_{\sigma \in S_{l_i}} \Theta(\sigma, v) &\geq q_i^{\mathcal{O}}(y, v)\left(\sum_{a \in [l_i]} \delta_{i,a}(\frac{\ln 2T}{\Delta^2} - 1) - \frac{1}{4}\right) \\ &\geq q_i^{\mathcal{O}}(y, v)\left(\sum_{a \in [l_i]} (\frac{\ln 2T}{\Delta} - 1) - \frac{1}{4}\right) \\ &= q_i^{\mathcal{O}}(y, v)\left(l_i \cdot \ln(2T) - l_i - \frac{1}{4}\right) \\ &\geq 0,\end{aligned}$$

where the last inequality follows from $T \geq 1$ and $q_i^{\mathcal{O}}(y, v) \geq 0$. Thus, inequality (C.11) is true, and it finishes the proof.

a

C.8 Proof of Theorem 5.5.2

Proof. Suppose without loss of generality, that the arms within each agent is sorted in decreasing order, *i.e.,* for any i, $\mu_{i,1} \geq \mu_{i,2} \geq \ldots \mu_{i,l_i}$. Assume further that the agents are also sorted in decreasing order with respect to their best arm, *i.e.,* $\mu_{1,1} \geq \mu_{2,1} \ldots \geq \mu_{n,1}$. For $i \in [n]$, define $\Delta_i = \mu_{1,1} - \mu_{i,1}$, and for each $a \in [l_i]$, let $\delta_{i,a} = \mu_{i,1} - \mu_{i,a}$. Let $\hat{\mu}_i(M)$ be agent i's empirical average reward at the end of round M.

Let I_t be the arm selected at round t, and r_t be the corresponding reward. We define an intra-agent regret $\textsc{Reg}_i(t)$ to be the regret occurred in Phase 2 until when agent i is selected exactly t times.

Precisely,

$$\textsc{Reg}_i(t) = t\mu_{i,1} - \sum_{\tau=1}^{t} r_\tau \mathbf{1}\{ I_\tau \in \mathcal{S}_i \}$$

Claim C.8.1. *For any agent i and for any $x \in [0,1]$, we have*

$$\Pr[\,|\hat{\mu}_i(M) - \mu_{i,1}| \geq x\,] \leq \frac{1}{Mx} \sum_{a=1}^{l_i} \Big(\frac{2L^2 \ln(2M)}{\Delta^2} + 1 \Big)$$

Proof of the claim.

$$\Pr[\,|\hat{\mu}_i(M) - \mu_{i,1}| \geq x\,] = \Pr[\,|M\hat{\mu}_i(M) - M\mu_{i,1}| \geq Mx\,] \tag{C.12}$$

We observe that $|M\hat{\mu}_i(M) - M\mu_{i,1}|$ is exactly agent i's internal regret until he is selected M times. Thus, we obtain

$$(C.12) = \Pr[\,\textsc{Reg}_i(t) \geq Mx\,] \leq \frac{\mathbb{E}[\,\textsc{Reg}_i(t)\,]}{Mx}. \qquad \text{(by Markov's inequality)}$$

By Lemma C.5.1, under the single agent setting when only agent i exists, if we run ETC with m as specified in the theorem statement, we have

$$\mathbb{E}[\,\textsc{Reg}_i(t)\,] \leq \sum_{a=1}^{l_i} \Big(\frac{2\delta_{i,a} L^2 \ln(2M)}{\Delta^2} + 1 \Big).$$

Thus we have

$$(C.12) \leq \frac{1}{Mx}\sum_{a=1}^{l_i}\Big(\frac{2\delta_{i,a}L^2\ln(2M)}{\Delta^2}+1\Big) \leq \frac{1}{Mx}\sum_{a=1}^{l_i}\Big(\frac{2L^2\ln(2M)}{\Delta^2}+1\Big),$$

and it completes the proof of the claim. a

Define $\texttt{Clean}_i$ be the event $\{\hat{\mu}_i(M) - \mu_{i,1} \geq \frac{1}{2}\Delta\}$. Then, by the claim above, we have

$$\Pr[\,\texttt{Clean}_i^c\,] \leq \frac{1}{M}\sum_{a=1}^{l_i}\Big(\frac{4L^2\ln(2M)}{\Delta^3}+1\Big) \leq \frac{1}{M}\sum_{a=1}^{l_i}\Big(\frac{4L^2\ln(2M)}{\Delta^3}+1\Big)$$

This implies that

$$\Pr[\,\cap_{i=1}^{n}\texttt{Clean}_i\,] \geq 1 - \frac{1}{M}\sum_{i=1}^{n}\sum_{a=1}^{l_i}\Big(\frac{4L^2\ln(2M)}{\Delta^3}+1\Big).$$

Note that whenever $\cap_{i=1}^{N}\texttt{Clean}_i$ happens, the following holds for any $i \in [n]$.

$$\begin{aligned}
\hat{\mu}_1(M) &\geq \mu_{1,1} - \frac{\Delta}{2} && \text{(Due to } \texttt{Clean}_1\text{)}\\
&\geq \mu_{i,1} + \frac{\Delta}{2} && \text{(Since } \mu_{1,1} - \mu_{i,1} \geq \Delta\text{)}\\
&\geq \hat{\mu}_i(M), && \text{(Due to } \texttt{Clean}_i\text{)}
\end{aligned}$$

and thus agent 1 will be selected in the Phase 1's exploration phase. In this case, the overall regret can

be decomposed as follows.

$$\begin{aligned}
\textsc{Reg}(T) &\leq (Mn + \mathbb{E}\,[\,\textsc{Reg}_i(T)\,]) + T \cdot \frac{1}{M} \sum_{i=1}^{n} \sum_{a=1}^{l_i} \Big(\frac{4L^2 \ln(2M)}{\Delta^3} + 1 \Big) \\
&\leq Mn + \sum_{a=1}^{l_i} \Big(\frac{2\delta_{i,a} L^2 \ln(2M)}{\Delta^2} + 1 \Big) + \frac{T}{M} \sum_{i=1}^{n} \sum_{a=1}^{l_i} \Big(\frac{4L^2 \ln(2M)}{\Delta^3} + 1 \Big) \\
&\leq Mn + \Big(\frac{2L^3 \ln(2M)}{\Delta^2} + l \Big) + \frac{T}{M} \Big(\frac{4nL^3 \ln(2M)}{\Delta^3} + nL \Big).
\end{aligned}$$

Note that $M = \sqrt{T \ln T}$ optimizes this upper bound (up to constant), and we finally obtain

$$\begin{aligned}
\textsc{Reg}(T) &\leq \sqrt{T \ln T} n + \frac{L^3 \ln(T \ln T)}{\Delta^2} + \frac{\sqrt{T}}{\sqrt{\ln T}} \Big(\frac{2nL^3 \ln(2T \ln T) + \Delta^3 nL}{\Delta^3} \Big) \\
&= O\Big(\frac{nL^3 \sqrt{T \ln T}}{\Delta^3} \Big),
\end{aligned}$$

and it finishes the proof. a

Appendix D: Optimal Contest beyond Convexity

D.1 Preliminaries for Section 4.4

D.1.1 Preliminaries on Topology

D.1.1.1 Basic of Topology

Topological space. A topology on X is a collection (topology) $\mathcal{T}$ of subsets of X satisfying the following:

1. $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$.
2. $\mathcal{T}$ is closed under any (possibly infinite) union of members of $\mathcal{T}$.
3. $\mathcal{T}$ is closed under any finite intersection of members of $\mathcal{T}$.

Open sets. The members of $\mathcal{T}$, *i.e.,* each $U \in \mathcal{T}$, are called *open sets*. A representative example is the standard topology with $X = \mathbb{R}$, where $\mathcal{T}$ is any union of open intervals in $\mathbb{R}$ with respect to one-dimensional Euclidean metric.

The pair $(X, \mathcal{T})$ is called a *topological space*, where we often write X as a topological space if underlying topology $\mathcal{T}$ is clear from the context. A subset $U \subset X$ is closed in $\mathcal{T}$ if $X \setminus U \in \mathcal{T}$.

Neighborhood, Hausdorff space, and continuous function. Given a topological space X and $x \in X$, a subset $V \subseteq X$ is a neighborhood of x if there exists an open set $U \in \mathcal{T}$ such that $x \in U$ and $U \subset V$. We say two points $x, y \in X$ can be separated by neighborhoods if there exists a

neighborhood U of x and V of y such that $U \cap V = \emptyset$. X is *Hausdorff space* if any two distinct points $x, y \in X$ are separated by neighborhoods.

A function $f : X \to Y$ between topological spaces is called *continuous* if for every point $x \in X$, and for every neighborhood V of $f(x)$ in Y, there exists a neighborhood U of x in X such that $f(U) \subseteq V$.

In our problem setup, the action space X is an Euclidean metric space, which is a well-known to be a Hausdorff space [230]. Looking forward, we will use the existential results with a so-called Hausdorff game (to be defined) by [88], which requires the action space to be the Hausdorff space.

Basis of topology. Given a topological space $(X, \mathcal{T})$, a collection $\mathcal{B}$ of open sets from $\mathcal{T}$ is called a basis of $\mathcal{T}$ if every open set $U \in \mathcal{T}$ can be expressed as a union of elements from $\mathcal{B}$. Equivalently, for any open set $U \in \mathcal{T}$ and any point $x \in U$, there exists $B \in \mathcal{B}$ such that $x \in B \subseteq U$.

For a natural number n, let X_i be a topological space for $i \in [n]$. Let $X := \prod_{i \in [n]} X_i$. The product topology on X is the topology whose basis consists of all sets of the form $U_1 \times U_2 \times \ldots \times U_n$, where U_i is an open set in X_i for $i \in [n]$.

D.1.1.2 Dual Space and Weak* Topology

Our analysis relies on the *metrizability* of the space of mixed strategies, viewed as a topological space equipped with the Lévy-Prokhorov metric (to be defined shortly). Intuitively, it can be technically challenging to reason directly about open neighborhoods of mixed strategies (i.e., probability measures over the pure strategy space) using only the abstract topology. A more tractable approach is to metrize this topology using a suitable metric—such as the Lévy-Prokhorov metric—which induces a natural notion of distance derived from the topology on the pure strategy space. This metric-based

perspective enables a more concrete analysis of convergence and continuity in the space of mixed strategies.

To formalize this, we begin with the classical fact that if X is a compact Hausdorff space, then by the Riesz Representation Theorem, every positive bounded linear functional on $L(X)$ can be uniquely represented by a finite Borel measure on X, where the space $L(X)$ is defined as:

$$L(X) = \{f : X \to \mathbb{R} \mid f \text{ is continuous and linear}\}. \tag{D.1}$$

Then, $L(X)$ is a normed vector space[1] under the supremum norm:

$$\|f\| = \sup_{x \in X} |f(x)|.$$

Since X is compact, this norm is finite, and $L(X)$ is complete—making it a Banach space (see, e.g., [231]).

We define the dual space $L(X)^*$ as the set of all real-valued continuous linear functionals on $L(X)$, that is,

$$L(X)^* = \{\phi : L(X) \to \mathbb{R} \big| \forall f, g \in L(X), \forall \alpha, \beta \in \mathbb{R}, \phi(\alpha f + \beta g) = \alpha\phi(f) + \beta\phi(G) \text{ and } \phi \text{ is continuous}\}$$

The dual norm on $L(X)^*$ is given by

$$\|\phi\| = \sup\{|\phi(f)| : f \in L(X), \|f\| \leq 1\}, \qquad \phi \in L(X)^*.$$

[1] A normed vector space is a vector space on which a norm is defined.

In general, a dual space of of a normed vector space X is the set of all continuous linear functionals from X to $\mathbb{R}$ similarly defined as above.

Dual space and the space of probability measure. By the Riesz representation theorem, any $\phi \in L(X)^*$ corresponds to integration against a signed regular Borel measure on X, where the Riesz representation theorem can be written as follows:

Theorem D.1.1 (Riesz representation theorem, [231])**.** *Let X be a compact Hausdorff space and $L(X)$ be the real vector space of all real-valued continuous bounded linear functionals on X, as defined in Equation (D.1). If $\phi \in L(X)^*$ is a positive linear functional such that $\phi(f) \geq 0$ for all $f \in L(X)$ with $f(x) \geq 0$ for every $x \in X$ and $||\phi|| = 1$ with the supremum norm, then there exists a unique Borel probability measure μ on X such that*

$$\phi(f) = \int_X f d\mu, \qquad \forall f \in L(X).$$

Let M denote the set of (Borel) probability measures over a compact Hausdorff space X, *i.e.,* the space of mixed strategies. Then, the Riesz representation theorem asserts the one-to-one correspondence[2] between the space of positive linear functional ϕ on $L(X)$ with $\|\phi\| = 1$ and M, *i.e.,* the space of mixed strategy. Such equivalence will be helpful in endowing the space of mixed strategies with a proper topology:

Definition D.1.2 (Weak* topology)**.** Let X be a normed vector space over $\mathbb{R}$ and X^* be its dual space, *i.e.,* the space of all linear continuous functionals $f : X \to \mathbb{R}$. For each $x \in X$, define a mapping $g_x : X^* \to \mathbb{R}$ as follows: $g_x(f) = f(x)$ for every $f \in X^*$. The weak* topology on X^* is defined as

[2]Note that the reverse direction is immediate as given a Borel probability measure μ on X, once we write $\phi(f) = \int_X f d\mu$ for $f \in L(X)$, this is a positive linear functional on $L(X)^*$.

the coarsest topology on X^* (with the fewest number of open sets) such that all mappings g_x for every $x \in X$ are continuous.

A seminal consequence of the weak* topology is that a sequence of functionals (f_i) in X^* converges to a functional $f \in X^*$ in the weak* topology if $f_i(x)$ converges to $f(x)$ for every $x \in X$ in a pointwise manner [230].

Throughout we consider the set of probability measures endowed with the weak* topology.

Metrizability. A topological space $(X, \mathcal{T})$ is called *metrizable* if its topology $\mathcal{T}$ is induced by a metric. Formally, if there exists a metric d on the set such that the open sets defined by d[3] If X is a metric space, it is known that $L(X)^*$ is metrizable with a proper metric, stated as follows.

Theorem D.1.3 ([232]). *If X is a metric space and M is a bounded subset of its dual space $L(X)^*$ (the space of bounded linear continuous functionals), then M, when endowed with the weak* topology, is a metrizable topological space. That is, it is homeomorphic*[4] *to a metric space under the Lévy-Prokhorov metric induced by the metric on X.*

Note that the Lévy-Prokhorov metric is defined as follows.

Definition D.1.4 (Lévy-Prokhorov metric)**.** Let (X, d) be a metric space given a distance function (metric) d with its Borel sigma algebra $\mathcal{B}(X)$. Let $\mathcal{P}(X)$ denote the collection of all Borel probability measures on the measurable space $(M, \mathcal{B}(X))$. For a subset $A \subseteq X$, define the ε-neighborhood of A by

$$A^\varepsilon = \{p \in X : \exists q \in A, d(p, q) < \varepsilon\} = \cup_{p \in A} B_\varepsilon(p),$$

[3]Given a metric space (X, d), a subset $U \subseteq X$ is called open if for all $x \in U$, there exists $\varepsilon > 0$ such that $B_d(x, \varepsilon) \subseteq U$ where $B_d(x, \varepsilon) = \{y \in X : d(x, y) < \varepsilon\}$.

[4]A function $h : X \to Y$ between two topological spaces (with some topologies endowed) if it is a bijection, continuous, and also its inverse is continuous. If such a function exists, we say two topological spaces are homeomorphic.

where $B_\varepsilon(p)$ is the open ball of radius ε at p, *i.e.,* $B_\varepsilon(p) = \{q \in X : d(p,q) < \varepsilon\}$. Then, the Lévy-Prokhorov metric $\pi : \mathcal{P}(M)^2 \to [0,\infty]$ is defined by setting the distance between two probability measures μ and ν to be

$$\pi(\mu,\nu) := \inf \left\{ \epsilon > 0 : \begin{array}{l} \mu(A) \leq \nu(A^\epsilon) + \epsilon \\ \nu(A) \leq \mu(A^\epsilon) + \epsilon \end{array} \quad \text{for all } A \in \mathcal{B}(X) \right\}.$$

Intuitively, $\pi(\mu,\nu)$ is the smallest ϵ such that for every measurable set A, the measure μ of A is close to the measure ν of the ϵ-expanded version of A, and vice versa. This captures how much the two probability distributions differ in terms of how they spread mass over points that are close together in the metric d.

Thus, Theorem D.1.3 immediately asserts that Lévy-Prokhorov metric metrizes the weak* topology on the space of bounded linear functionals, which in fact is the space of mixed strategies by Theorem D.1.1, *i.e.,* any open ball with respect to the Lévy-Prokhorov metric is an open set in the weak* topology. The following proposition summarizes the preliminaries we require to prove the existential result.

Proposition D.1.5. *The family of mixed strategies X^* over an action space X, where X is a metric space, can be endowed with the weak* topology, which can be metrized by the Lévy–Prokhorov metric. In particular, if $\phi : X \to \mathbb{R}$ is bounded and continuous, then the map $h : \mu \mapsto \int \phi \, d\mu$ is continuous with respect to the weak* topology on X^*.*

D.1.2 Preliminaries on Games and Existential Results

Technically speaking, to argue the existence of a symmetric MNE, we need to analyze the expected payoff of mixed strategies that lie within an open neighborhood of a symmetric strategy in order to apply the argument by [88]. To this end, we rigorously define the notion of an open neighborhood around a mixed strategy (i.e., a probability measure on a compact set) using the weak* topology, as formalized in Definition D.1.2.

Since reasoning directly over the topological space can be challenging, we exploit the fact that the space of probability measures on a compact Hausdorff space is metrizable under the L'evy-Prokhorov metric. This enables us to treat the topological space as a metric space—namely, we analyze the topology of probability measures on X (endowed with the weak* topology) using a metric induced from the original topological (metric) space X.

For readers unfamiliar with basic topological concepts, we refer to Section D.1.1 for background material.

D.1.2.1 Game and Mixed Extension

A game $G = (X_i, u_i)_{i\in[n]}$ consists of n players, a set of pure strategy X_i and payoff functions $u_i : X \to \mathbb{R}$ where $X = \times_{i\in[n]} X_i$ for $i \in [n]$. We assume throughout that X_i is a compact Hausdorff space, and we say the game is *Hausdorff* game in such case. Throughout, the product of any number of sets is endowed with the product topology. Let M_i denote the set of (Borel) probability measures[5] on X_i for $i \in [n]$, *i.e.,* the set of mixed strategies over X_i.

Given a game $G = (X_i, u_i)_{i\in[n]}$ and the set of mixed strategies M_i defined above, we extend

[5] A Borel measure on a topological space is a measure that is defined on all open sets.

each u_i to have its domain as $M = \times_{i\in[n]} M_i$ by defining

$$u_i(\mu) = \int_X u_i(x) d\mu,$$

for all $\mu \in M$. We denote by $\bar{G} = (M_i, u_i)_{i\in[n]}$ the mixed extension of G.

Quasi-symmetric game. Consider a *game* $G = (X_i, u_i)_{i\in[n]}$ (which is possibly a mixed extension of another original game G'). We say that G is a *quasi-symmetric* game if $X = X_1 = \ldots X_n$ and for any $x, y \in X$, it follows that

$$u_1(x, y, \ldots, y) = u_2(y, x, y, \ldots, y) = \ldots = u_n(y, \ldots, y, x).$$

Define a quasi-symmetric game's *diagonal* payoff function $v : X \to \mathbb{R}$ as

$$v(x) = u_1(x, \ldots, x) = u_n(x, \ldots, x)$$

for every $x \in X$.

D.1.2.2 Better-reply Secureness

In proving the existence of a symmetric MNE, the notion of better-reply secure is crucial. To that end, we first introduce a notion of securing a payoff.

Definition D.1.6 (Secure a payoff)**.** Given a quasi-symmetric game $G = (X, u_i)_{i\in[n]}$, player i can

secure a payoff of $\alpha \in \mathbb{R}$ *along the diagonal* at $(x, \ldots, x)$ if there exists $\bar{x} \in X$ such that

$$u_i(x', \ldots, \bar{x}, \ldots, x') \geq \alpha$$

for every x' in some open neighborhood of $x \in X$.[6]

The following diagonally better-reply secureness guarantees that any player could secure a strictly larger payoff along the diagonal, given a symmetric strategy profile which is not an equilibrium.

Definition D.1.7 (Diagonally better-reply secure)**.** Given a quasi-symmetric game $G = (X, u_i)_{i\in[n]}$, let $(x, v(x)) \in X \times \mathbb{R}$ be the pair of an action $x \in X$ and diagonal payoff function $v(x)$. Then, G is *diagonally better-reply secure* if for any $(x, v(x))$ with $x \in X$ such that $(x, \ldots, x)$ is not an equilibrium, there exists a player i who can secure a payoff strictly above $v(x)$ along the diagonal at $(x, \ldots, x)$ per Definition D.1.6.

The seminal result by [88] prove that any quasi-symmetric game possesses a symmetric MNE under the diagonally better-reply secureness and semi-continuity of the payoff function.

Theorem D.1.8 (Corollary 5.3, [88])**.** *Suppose* $G = (X_i, u_i)_{i\in[n]}$ *is a quasi-symmetric game with compact Hausdorff action space. Then,* G *has a symmetric MNE if its mixed extension* $\bar{G}$ *is diagonally better-reply secure and each* $u_i(\mu, \ldots, \mu)$ *is upper semi-continuous over* μ*.*

[6]When dealing with G as a mixed extension of another original game, the open neighborhood refers to the open neighborhood with respect to the underlying topology over the space of mixed strategies.

D.2 Omitted Proofs in Section 4.4

D.2.1 Proof of Theorem 4.4.2

We begin with the following result, establishing that the induced game is both quasi-symmetric and a Hausdorff game (as defined in Section D.1.2).

Lemma D.2.1. *The induced game given a nontrivial policy is both quasi-symmetric and a Hausdorff game (as defined in Section D.1.2).*

Proof. The quasi-symmetry immediately follows from the fact that every producer is homogeneous in a sense that they have the same cost function and thereby the same payoff as a function over the quality. This is true even if the mixed strategy is not atomless as we assume uniform random tie-breaking rule.

To prove that the game is Hausdorff, it suffices to show that its action space is compact. Given any policy, for any symmetric MNE, each provider must receive nonnegative expected utility for every action q in the support of the MNE, since playing $q = 0$ yields utility zero.

On the other hand, if the support contains $q > 1$, the expected utility becomes strictly negative due to the cost function as $c(q) = q^\beta > 1$. Therefore, it is without loss of generality to restrict the action space to the interval $[0, 1]$, which is compact.

a

Next, we prove that the utility function of each producer, viewed as a function over mixed strategies, is upper semi-continuous at any symmetric strategy profile. As discussed in Section D.1.1, to handle continuity with respect to the mixed strategy μ—that is, a probability measure—we consider the space of mixed strategies endowed with the weak* topology. This topology is metrizable via

the L'evy-Prokhorov metric, as noted in Proposition D.1.5. We refer the reader to Section D.1.1 for additional background.

Lemma D.2.2. *For any nontrivial policy* $\boldsymbol{p}$*, at any symmetric mixed strategy profiles* $\boldsymbol{\mu} = (\mu, \dots, \mu)$

Proof. Due to the symmetric of the game, for any symmetric mixed strategy, it is straightforward to see that the overall expected revenue of each producer is simply $1/n$. Thus, the expected utility can be written as

$$\mathcal{U}_i(\mu; \bar{\mu}_{-i}) = \frac{1}{n} - \int_0^1 c(x) d\mu.$$

As we endow the space of set of mixed strategies μ with the weak* topology, the cost function $c(x)$ is continuous, and thus $\int_0^1 c(x)d\mu$ is continuous in μ by Proposition D.1.5. This concludes that the utility function $\mathcal{U}_i$ is upper semi-continuous with respect to μ as continuity implies upper semi-continuity..

a

Now, for ease of exposition, we define

$$R(q; \boldsymbol{\mu}_{-i}) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1 - F(q))^{i-1}, \tag{D.2}$$

which represents the expected revenue from choosing quality level q when all other producers follow the mixed strategy described by F. That is, as noted in Equation (D.4),

$$\mathcal{U}_i(q; \boldsymbol{\mu}_{-i}) = R(q; \boldsymbol{\mu}_{-i}) - c(q),$$

as presented in Lemma 4.4.4, under the assumption that F has no point masses. We often write $R(q)$

when the strategies of the other producers are clear from context.

The following lemma will be useful in proving diagonally better-reply secureness.

Lemma D.2.3. *For any nontrivial policy $\boldsymbol{p}$, given the other producers' symmetric mixed strategies $\boldsymbol{\mu}_{-i}$ with corresponding c.d.f. F, the expected revenue $R(q; \boldsymbol{\mu}_{-i})$ is monotone increasing in q.*

Proof. Fix q, and consider a random variable $Y_q \sim \text{Binomial}(n-1, 1-F(q))$. Then

$$\Pr\left[\, Y_q = i-1 \,\right] = \binom{n-1}{i-1} F(q)^{n-i} (1-F(q))^{i-1}.$$

Therefore, by Equation (D.2), we have

$$\begin{aligned} R(q; \boldsymbol{\mu}_{-i}) &= \sum_{i=1}^{n} p_i \Pr\left[\, Y_q = i-1 \,\right] \\ &= \sum_{i=1}^{n} p_i \Pr\left[\, i = Y_q + 1 \,\right] \\ &= \mathbb{E}\left[\, p_{Y_q+1} \,\right]. \end{aligned}$$

Now if $q \geq q'$, then $F(q) \geq F(q')$, and thus Y_q is stochastically dominated by $Y_{q'}$. Hence, we have $\mathbb{E}\left[\, f(Y_q) \,\right] \geq \mathbb{E}\left[\, f(Y_{q'}) \,\right]$ for any non-increasing function f, *e.g.,* see [233]. Since p_{Y+1} is non-increasing in Y as $p_1 \geq p_2 \geq \ldots \geq p_n$, we have $\mathbb{E}\left[\, p_{Y_q+1} \,\right] \geq \mathbb{E}\left[\, p_{Y_{q'}+1} \,\right]$, which finishes the proof. a

To prove the existence of the symmetric MNE, the final step is to prove the diagonally better-reply secureness by [88]; see Theorem D.1.8.

Lemma D.2.4. *For any policy $\boldsymbol{p}$ with $p_1 \geq p_2 \geq \ldots \geq p_n$, the induced game is diagonally better-reply secure per Definition D.1.7.*

Proof. Let μ be a mixed strategy that does *not* constitute a symmetric MNE, and let $u = \mathcal{U}(\mu; \mu, \ldots, \mu)$ denote the corresponding diagonal payoff. Per Definition D.1.7, to establish diagonal better-reply security, it suffices to find another strategy μ' such that $\mathcal{U}(\mu'; \mu'', \ldots, \mu'') > u$ for any μ'' in an open neighborhood $B_\delta(\mu)$, for some $\delta > 0$, with respect to the Lévy-Prokhorov metric. Here, $B_\delta(\mu) = \{\nu : d(\mu, \nu) < \delta\}$ where $d(\cdot, \cdot)$ denotes the Lévy-Prokhorov metric. Since μ is not an equilibrium, there exists a pure strategy $q^{\text{dev}} \in \mathbb{R}$ such that

$$\mathcal{U}(q^{\text{dev}}; \mu, \ldots, \mu) > \mathcal{U}(\mu; \mu, \ldots, \mu) = u.$$

To see why q^{dev} must exist, suppose for the sake of contradiction that no pure strategy strictly improves the payoff. Then, for any mixed strategy ν, we have

$$\mathcal{U}(\nu; \mu, \ldots, \mu) = \int_{q \in \text{supp}(\nu)} \mathcal{U}(q; \mu, \ldots, \mu)\, d\nu(q) \leq \int_{q \in \text{supp}(\nu)} u\, d\nu(q) = u,$$

which implies that no mixed strategy can strictly increase the payoff—a contradiction. Therefore, such a pure strategy q^{dev} must exist.

To construct a secure deviation, we perturb q^{dev} to a nearby pure strategy q^{sec} such that μ does not have an atom at q^{sec} and

$$\mathcal{U}(q^{\text{sec}}; \mu, \ldots, \mu) > u.$$

This is possible since any probability distribution has at most countably many atoms. Let F be the c.d.f. corresponding to μ. Since q^{sec} does not have an atom, by (D.2)

$$R(q^{\text{sec}}; \mu, \ldots, \mu) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q^{\text{sec}})^{n-i} (1 - F(q^{\text{sec}}))^{i-1},$$

and thus

$$\mathcal{U}(q^{\text{sec}}; \mu, \dots, \mu) = R(q^{\text{sec}}; \mu, \dots, \mu) - c(q^{\text{sec}}) > u.$$

Due to the continuity of F, we can find a sufficiently small $\varepsilon > 0$ such that μ does not have an atom at $q^{\text{sec}} - \varepsilon$ and satisfies

$$\sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q^{\text{sec}} - \varepsilon)^{n-i} (1 - F(q^{\text{sec}} - \varepsilon))^{i-1} - c(q^{\text{sec}}) > u. \tag{D.3}$$

This holds because as $\varepsilon \to 0$, the left-hand side converges to $\mathcal{U}(q^{\text{sec}}; \mu, \dots, \mu) > u$. Moreover, since any distribution has at most countably many atoms, there must exist such a sufficiently small ε for which μ has no atom in the interval $[q^{\text{sec}} - \varepsilon, q^{\text{sec}}]$. Note that such ε can be found arbitrarily small since if we consider any small neighborhood $[q^{sec} - \varepsilon', q^{sec}]$ of q^{sec}, there exists at most countably many atoms on it, and thus we can find sufficiently small ε with no atom on it. We will prove that q^{sec} is the desired strategy that secures payoff larger than u along the diagonal at $\bar{\mu}$.

Let $A := \{q' : q' < q^{\text{sec}}\}$ and $A^{\varepsilon} := \{q' : q' < q^{\text{sec}} - \varepsilon\}$. Then $F(q^{\text{sec}}) = \mu(A)$ and $F(q^{\text{sec}} - \varepsilon) = \mu(A^{\varepsilon})$. Rewriting (D.3), we get

$$\Gamma := \sum_{i=1}^{n} p_i \binom{n-1}{i-1} \mu(A^{\varepsilon})^{n-i} (1 - \mu(A^{\varepsilon}))^{i-1} - c(q^{\text{sec}}) > u.$$

Let $A' := \{q' : \exists q \in A^{\varepsilon} \text{ s.t. } |q' - q| < \varepsilon\}$. By the definition of the Lévy-Prokhorov metric (see Definition D.1.4), for any $\mu' \in B_{\varepsilon}(\mu)$,

$$\mu(A^{\varepsilon}) \leq \mu'(A') + \varepsilon.$$

Since any $q' \in A'$ satisfies $q' < q^{\text{sec}}$, we also have

$$\mu'(A) \geq \mu'(A') \geq \mu(A^\varepsilon) - \varepsilon.$$

Therefore, for any $\mu' \in B_\varepsilon(\mu)$, the payoff satisfies:

$$\begin{aligned}
\mathcal{U}(q^{\text{sec}}; \mu', \ldots, \mu') &\geq \sum_{i=1}^{n} p_i \binom{n-1}{i-1} \mu'(A)^{n-i}(1-\mu'(A))^{i-1} - c(q^{\text{sec}}) \\
&\geq \sum_{i=1}^{n} p_i \binom{n-1}{i-1} (\mu(A^\varepsilon) - \varepsilon)^{n-i}(1-\mu(A^\varepsilon)+\varepsilon)^{i-1} - c(q^{\text{sec}}) \\
&= \Gamma + \Delta,
\end{aligned}$$

where the second inequality follows from the monotonicity in Lemma D.2.3 and

$$\Delta := \sum_{i=1}^{n} p_i \binom{n-1}{i-1} \left[(\mu(A^\varepsilon) - \varepsilon)^{n-i}(1-\mu(A^\varepsilon)+\varepsilon)^{i-1} - \mu(A^\varepsilon)^{n-i}(1-\mu(A^\varepsilon))^{i-1} \right].$$

Since $\Delta \to 0$ as $\varepsilon \to 0$, we can choose ε small enough so that $\Delta > -(\Gamma - u)$, implying

$$\mathcal{U}(q^{\text{sec}}; \mu', \ldots, \mu') > u.$$

This concludes the proof.

a

Proof of Theorem 4.4.2. Finally, the proof immediately follows from Theorem D.1.8 combined with lemmas D.2.1, D.2.3, D.2.2, and D.2.4. a

D.2.2 Proof of Proposition 4.4.1

Proof. Given a policy $\mathbf{p}$, suppose there exists a PNE $q_1 \geq q_2 \geq \ldots \geq q_n$. If $q_i \neq 0$ but $p_i = 0$, producer i can effectively deviates to $q_i = 0$ while decreasing the cost, which is a contradiction. Thus, if $q_i \neq 0$ then we always have $p_i \neq 0$. Let i be the largest index where $q_i \neq 0$. Suppose there exists no tie among producers $j \in [i-1]$, *i.e.,* $q_1 > q_2 > \ldots > q_i$. Any producer $j \in [i]$ can deviate to $q_j(1-\varepsilon)$ for sufficiently small $\varepsilon > 0$ while respecting $q_j(1-\varepsilon) > q_{j+1}$, which only decreases the cost. Thus, this cannot be an equilibrium. Suppose now that there exists a tie at j such that $q_j = q_{j+1}$. Suppose there exists $k \geq 2$ such agents where j be the smallest index, *i.e.,* $q_j = q_{j+1} = \ldots = q_{j+k-1}$. Due to the uniformly random tie-breaking rule, all the producers from j to $j+k-1$ obtain revenue of $(p_j + p_{j+1} + \ldots p_{j+k-1})/k$. If $p_j = p_{j+k-1}$, then producer j can slightly decrease the quality while still receiving the revenue of $p_{j+k-1} = (p_j + p_{j+1} + \ldots p_{j+k-1})/k$ and decreasing the cost. Thus, $p_j \neq p_{j+k-1}$. In this case, producer j can instead increase the quality from q_j to $q_j + \varepsilon$ and solely obtain $p_j > (p_j + p_{j+1} + \ldots p_{j+k-1})/k$ while only increasing cost function sufficiently small so that the overall utility increases due to the continuity of the payoff function. This finishes the proof. a

D.2.3 Proof of Proposition 4.4.3

Proof. To prove that the strategy is atomless, fix an agent i. Assume, for the sake of contradiction, that there exists a point p such that agent i plays p with positive probability $\alpha > 0$. Let $p' = p + \varepsilon$ for some $\varepsilon > 0$. Let E be the event that all other producers $j \in [n] \setminus \{i\}$ also play p, *i.e.,* , $\Pr[E] = \alpha^{n-1}$.

Conditioned on E, playing p yields expected revenue $1/n$. On the other hand, playing p' yields expected revenue at least $p_1 > 1/n$ for any $\varepsilon > 0$, since $\mathbf{p}$ is nontrivial. Thus, by the continuity of the cost function, agent i can deviate to $p + \varepsilon$ for sufficiently small ε and strictly increase their utility under

the event E.

Since this deviation does not decrease the revenue under E^c, one can find sufficiently small ε to strictly improve agent i's utility overall—a contradiction.

To prove that the distribution function F is strictly increasing, suppose that $(a, b) \subseteq [q_{\min}, q_{\max}]$ is the subinterval of maximal length such that F remains constant on (a, b), where

$$q_{\max} = \sup\{y : F(y) < 1\}.$$

Note that $q_{\min} \geq 0$ and $q_{\max} \leq 1$ because playing quality 1 yields nonpositive utility.

Clearly, any agent placing mass on b can move that mass to a, thereby strictly increasing their utility—a contradiction. Hence, F must be strictly increasing on its support. Similarly, if $a > 0$, any agent can move its mass from a to 0 to strictly increase their utility, which is a contradiction. So, F is strictly increasing on $[0, q_{\max}]$. a

D.2.4 Proof of Lemma 4.4.4

Proof. From Proposition 4.4.3, we know that μ is atomless, *i.e.,* tie-breaking happens with probability 0. Thus, we can write the producer i's payoff as follows:

$$\begin{aligned}
\mathcal{U}_i(q, \boldsymbol{\mu}_{-i}, \mathbf{p}) &= \mathcal{R}_i(q, \boldsymbol{\mu}_{-i}, \mathbf{p}) - c(q) \\
&= \mathbb{E}\left[\left(\sum_{i\in[n]} \mathbf{1}\left\{\, I = i \,\right\}\right) - c(q)\right] \\
&= \left(\sum_{i\in[n]} \Pr\left[\,\text{Provider } i\text{'s content is recommended to user } i\,\right]\right) - c(q) \\
&= \left(\sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1}\right) - c(q). \qquad \text{(D.4)}
\end{aligned}$$

For the user welfare, let E_i be the random event such that the chosen content has the i-th highest quality among $[n]$ and I be the variable that indicates the index of the content chosen by the RS. Expanding using the definition of $\mathbf{p}$, we obtain:

$$\begin{aligned}
\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) &= \mathbb{E}_{\mathbf{q}\sim\bar{\mu}}\left[\sum_{i=1}^{n} \mathbf{1}\left\{\, I = i \,\right\} \cdot q_i\right] \\
&= \mathbb{E}_{\mathbf{q}\sim\bar{\mu}}\left[\sum_{i=1}^{n} \mathbf{1}\left\{\, E_i \,\right\} \cdot q_{(i)}\right] \\
&= \sum_{i=1}^{n} p_i \mathbb{E}_{\mathbf{q}\sim\bar{\mu}}\left[q_{(i)}\right].
\end{aligned}$$

Finally, the equation for the platform quality immediately follows from the fact that $\bar{\mu}$ is symmetric strategy profiles.

a

D.2.5 Proof of Theorem 4.4.5

Proof. Recall first that

$$a_i(x) = \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}$$
$$h(x,p) = \sum_{i=1}^{n} p_i\, a_i(x) \quad (x \in [0,1]).$$

By our existence result for symmetric MNE, a symmetric equilibrium μ exists for every nontrivial policy p. By (D.12), we have that

$$u_0 + q^\beta \;=\; \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{\,n-i}\big(1-F(q)\big)^{i-1} \;=\; h\big(F(q),p\big) \quad , \tag{D.5}$$

for all $q \in [q_{\min}, q_{\max}]$ where $q_{\min}$ and $q_{\max}$ denote minimal and maximal value in the support of F. Note that Proposition 4.4.3 yields $q_{\min} = 0$. Thus, plugging $q = 0$ into (D.5) and using $F(0) = 0$ gives $u_0 = p_n$.

Recall that the c.d.f. F is atomless and strictly increasing on some interval $[0, q_{\max}]$ by Proposition 4.4.3. Hence $h(\cdot, p)$ has a unique continuous inverse on $[h(0,p), h(1,p)] = [p_n, p_1]$.

Thus, we have

$$F(q) \;=\; h^{-1}\big(p_n + q^\beta\,; p\big), \tag{D.6}$$

for all $q \in [0, q_{\max}]$. Notice that this determines the c.d.f. F uniquely once its support is given. Since $F(0) = 0$ we must have $h(0,p) = p_n$, and since $F(q_{\max}) = 1$ we must have $p_n + q_{\max}^\beta = h(1,p) = p_1$, i.e., $q_{\max} = (p_1 - p_n)^{1/\beta}$.

Let us confirm again that the support is always $[0, q_{\max}]$ for $q_{\max} = (p_1 - p_n)^{1/\beta}$. For any $q > q_{\max}$, all other players' draws lie in $[0, q_{\max}]$, so a deviation to q surely earns the top rank and revenue p_1 due to atomlessness, while incurring cost q^β. Since $q^\beta \geq q_{\max}^\beta = p_1 - p_n$, the deviator's payoff $p_1 - q^\beta \leq p_n = u_0$, hence there is no profitable deviation above the support. If the minimum value in the support of F is $q_{\min} > 0$, then any play at $q_{\min}$ can deviate to 0 by decreasing its cost and still ensure revenue p_n. Thus, any symmetric equilibrium has the distribution F on the support $[0, q_{\max}]$ for $q_{\max} = (p_1 - p_n)^{1/\beta}$.

Finally, to conclude, suppose $\tilde{F}$ be a c.d.f. of another symmetric MNE. Notice that it needs to satisfy (D.6) on the support $[0, q_{\max}]$. Since $h(\cdot, p)$ is strictly increasing by Lemma D.5.1, (D.6) implies that $\tilde{F}(q) = h^{-1}(p_n + q^\beta; p) = F(q)$ for all q in the support. Thus the symmetric MNE is unique. a

D.3 Omitted Proofs in Section 4.5

D.3.1 Proof of Theorem 4.1.1

Before proving the results, we introduce further notations:

$$\psi(x, \mathbf{p}) = \sum_{i=2}^{n-1} p_i (n-1) \left(\binom{n-2}{i-1} x^{n-i-1}(1-x)^{i-1} - \binom{n-2}{i-2} x^{n-i}(1-x)^{i-2} \right)$$
$$+ p_1(n-1)x^{n-2} - p_n(n-1)(1-x)^{n-2}, \tag{D.7}$$
$$\phi(x, \mathbf{p}) = \sum_{i=1}^{n} \left(1 - \sum_{j=1}^{i} \binom{n}{j-1} (1-x)^{j-1}(x)^{n-j+1}) \right) p_i \tag{D.8}$$
$$= 1 - \sum_{i=1}^{m} \binom{n}{i-1} (1-x)^{i-1} x^{n-i+1} (\sum_{j=i}^{n} p_j), \tag{D.9}$$

Furthermore, we require several technical lemmas. We begin by establishing connections be-

tween the notations defined above and those introduced in Section 4.5.

The following is a well-known characterization of the c.d.f. of order statistics for identically and independently distributed (i.i.d.) random variables.

Lemma D.3.1 ([234]). *Given n i.i.d. random variables $X_1, .., X_n$, let F be its c.d.f. and $F_{(i)}$ be the c.d.f. of the i-th largest order statistic. Then, for any $i \in [n]$,*

$$F_{(i)}(x) = \sum_{j=1}^{i} \binom{n}{j-1} (1 - F(x))^{j-1} F(x)^{n-j+1}.$$

We omit the proof, which is routine and covered in standard texts.

Lemma D.3.2. *The function $\psi(x, \boldsymbol{p})$, defined in Equation (D.7), represents the derivative of $h(x, \boldsymbol{p})$ with respect to x, where $h(x, \boldsymbol{p}) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}$ is given in Equation (4.1).*

Proof. Note that

$$\frac{dh(x, \mathbf{p})}{dx} = \frac{d \sum_{i=1}^{n} p_i a_i(x)}{dx}. \tag{D.10}$$

Recall that $a_i(x) = \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}$. If $i = 2, \ldots, n-1$, we have

$$\begin{aligned}
\frac{da_i(x)}{dx} &= \binom{n-1}{i-1}(n-i)x^{n-i-1}(1-x)^{i-1} - \binom{n-1}{i-1}(i-1)x^{n-i}(1-x)^{i-2} \\
&= \frac{(n-1)!}{(i-1)!(n-i-1)!} x^{n-i-1}(1-x)^{i-1} - \frac{(n-1)!}{(i-2)!(n-i)!} x^{n-i}(1-x)^{i-2} \\
&= (n-1)\left(\binom{n-2}{i-1} x^{n-i-1}(1-x)^{i-1} - \binom{n-2}{i-2} x^{n-i}(1-x)^{i-2} \right).
\end{aligned}$$

For $a_1(x)$ and $a_n(x)$, we have

$$\frac{da_1(x)}{dx} = (n-1)x^{n-2}\,, \qquad \frac{da_n(x)}{dx} = -(n-1)(1-x)^{n-2}.$$

Plugging into (D.10) finishes the proof. a

The following is a well-known property of Bernstein polynomials, whose proof is provided for completeness.

Lemma D.3.3 ([235]). $\int_0^1 a_i(x)dx = 1/n$ *for any* $i \in [n]$, *where* $a_i(x)$ *is defined in Equation (4.1).*

Proof. Substitute the definition of $a_i(x)$:

$$\begin{aligned}\int_0^1 a_i(x)dx &= \int_0^1 \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}dx \\ &= \binom{n-1}{i-1}\int_0^1 x^{n-i}(1-x)^{i-1}dx\end{aligned}$$

The integral part is a Beta function, which is defined as:

$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}dt = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},$$

where $\Gamma(k) = \int_0 t^{k-1}e^{-t}dt$ which is $(k-1)!$ for positive integer k. Thus for positive integers p and q, this can be written as $B(p,q) = \frac{(p-1)!(q-1)!}{(p+q-1)!}$.

Comparing the integral $\int_0^1 x^{n-i}(1-x)^{i-1}dx$ with the Beta function definition, notice that the integral part can be written as $B(n-i+1,i)$, which can be written as:

$$B(n-i+1,i) = \frac{\Gamma(n-i+1)\Gamma(i)}{\Gamma((n-i+1)+i)} = \frac{(n-i)!(i-1)!}{(n)!}$$

Now, we also expand the binomial coefficient $\binom{n-1}{i-1}$:

$$\binom{n-1}{i-1} = \frac{(n-1)!}{(i-1)!(n-1-(i-1))!} = \frac{(n-1)!}{(i-1)!(n-i)!}$$

Substitute these expressions back into the integral:

$$\begin{aligned}\int_0^1 a_i(x)dx &= \frac{(n-1)!}{(i-1)!(n-i)!} \cdot \frac{(n-i)!(i-1)!}{n!} \\ &= \frac{(n-1)!}{n!} \\ &= \frac{(n-1)!}{n \cdot (n-1)!} \\ &= \frac{1}{n},\end{aligned}$$

which finishes the proof. a

Finally, we require the following relation between $\phi(x, \mathbf{p})$ and $h(x, \mathbf{p})$.

Lemma D.3.4. *It follows that*

$$\frac{d\phi(x,\boldsymbol{p})}{dx} = -nh(x,\boldsymbol{p}),$$

where $\phi(x,\boldsymbol{p})$ is defined in Equation (D.9) and $h(x,\boldsymbol{p})$ is defined in Equation (4.1).

Proof. We have

$$\begin{aligned}\frac{d\phi(x,\mathbf{p})}{dx} = -nx^{n-1} &+ \underbrace{\sum_{i=2}^{n} \binom{n}{i-1}(i-1)(1-x)^{i-2}x^{n-i+1}\sum_{j=i}^{n} p_j}_{(A)} \\ &- \underbrace{\sum_{i=2}^{n} \binom{n}{i-1}(n-i+1)(1-x)^{i-1}x^{n-i}\sum_{j=i}^{n} p_j}_{(B)}.\end{aligned}$$

Note that

$$\begin{aligned}(A) &= \sum_{i=2}^{n} \frac{n!}{(i-1)!(n-i+1)!}(i-1)(1-x)^{i-2}x^{n-i+1}\sum_{j=i}^{n} p_j \\ &= \sum_{i=2}^{n} \frac{n!}{(i-2)!(n-i+1)!}(1-x)^{i-2}x^{n-i+1}\sum_{j=i}^{n} p_j \\ &= n\sum_{i=2}^{n} \frac{(n-1)!}{(i-2)!(n-i+1)!}(1-x)^{i-2}x^{n-i+1}\sum_{j=i}^{n} p_j \\ &= n\sum_{i=1}^{n-1} \frac{(n-1)!}{(i-1)!(n-i)!}(1-x)^{i-1}x^{n-i}\sum_{j=i+1}^{n} p_j\end{aligned}$$

On the other hand,

$$\begin{aligned}(B) &= \sum_{i=2}^{n} \frac{n!}{(i-1)!(n-i+1)!}(n-i+1)(1-x)^{i-1}x^{n-i}\sum_{j=i}^{n} p_j \\ &= \sum_{i=2}^{n} \frac{n!}{(i-1)!(n-i)!}(1-x)^{i-1}x^{n-i}\sum_{j=i}^{n} p_j \\ &= n\sum_{i=2}^{n} \frac{(n-1)!}{(i-1)!(n-i)!}(1-x)^{i-1}x^{n-i}\sum_{j=i}^{n} p_j.\end{aligned}$$

Hence, we have

$$\begin{aligned}(A)-(B) &= nx^{n-1}\sum_{j=2}^{n} p_j - n(1-x)^{n-1}p_n + n\sum_{i=2}^{n-1} \frac{(n-1)!}{(i-1)!(n-i)!}(1-x)^{i-1}x^{n-i}\left(\sum_{j=i+1}^{n} p_j - \sum_{j=i}^{n} p_j\right) \\ &= nx^{n-1}\sum_{j=2}^{n} p_j - n(1-x)^{n-1}p_n - n\sum_{i=2}^{n-1} \frac{(n-1)!}{(i-1)!(n-i)!}(1-x)^{i-1}x^{n-i}p_i \\ &= nx^{n-1}\sum_{j=1}^{n} p_j - nx^{n-1}p_1 - n(1-x)^{n-1}p_n - n\sum_{i=2}^{n-1} \frac{(n-1)!}{(i-1)!(n-i)!}(1-x)^{i-1}x^{n-i}p_i \\ &= nx^{n-1}\sum_{i=1}^{n} p_i - n\sum_{i=1}^{n} \frac{(n-1)!}{(i-1)!(n-i)!}(1-x)^{i-1}x^{n-i}p_i\end{aligned}$$

Thus, we have

$$(A) - (B) = nx^{n-1} - n\sum_{i=1}^{n} a_i(x)p_i,$$

concluding that

$$\frac{d\phi(x,\mathbf{p})}{dx} = -n\sum_{i=1}^{n} a_i(x)p_i = -nh(x,\mathbf{p})\,. \tag{D.11}$$

a

We finally need the following correlation type inequality.

Lemma D.3.5 ([212]). *Let $\nu_1, \ldots, \nu_n$ be a probability measure on $\mathbb{R}$. If functions $f, g : \mathbb{R}^n \mapsto [0,1]$ are monotone increasing functions, then $\mathbb{E}_\nu[f \cdot g] \geq \mathbb{E}_\nu[f] \cdot \mathbb{E}_\nu[g]$, where ν is a product measure over $\nu_1, \ldots, \nu_n$. The same inequality holds when both f, g are monotone decreasing functions.*

Now we can proceed to the proof of the main theorem.

Proof of Theorem 4.1.1. Let $\boldsymbol{\mu}$ be the symmetric MNE and F be corresponding c.d.f. of the induced game given a nontrivial policy $\mathbf{p}$. For $\boldsymbol{\mu}$ to be the equilibrium, the following condition must hold:

$$u_0 + c(q) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1},$$

for some $u_0 \geq 0$ and for any $q \in \mathrm{supp}(F)$, as there exists no atom in $\bar{\mu}$ by Proposition 4.4.3. Note that u_0 is the expected payoff at the equilibrium.

Let $q_{(i)}$ denote the i-th largest quality among $(q_i)_{i\in[n]}$ where q_i is random quality realized for each agent $i \in [n]$ with respect to the mixed strategy μ. Let $F_{(i)}$ be the c.d.f. of $q_{(i)}$. Using Lemma 4.4.4,

the RS's problem to maximize the objective function can be written as the following optimization problem:

$$
\begin{aligned}
(\text{OPT}_1) \quad \underset{\mathbf{p}\in\Delta_m}{\text{maximize}} \quad & \alpha \sum_{i=1}^{n} p_i \mathbb{E}\left[\, q_{(i)} \,\right] + (1-\alpha)\mathbb{E}\left[\, q_i \,\right] \\
\text{subject to} \quad & u_0 + c(q) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1} \quad \forall q \in \operatorname{supp}(F) \\
& p_1 \geq p_2 \geq \ldots \geq p_n, \ \exists i, j : p_i \neq p_j \\
& \sum_{i=1}^{n} p_i = 1 \\
& F \text{ is c.d.f. of symmetric MNE given } \mathbf{p}
\end{aligned}
$$

Recall that $0 \in \operatorname{supp}(F)$ by Proposition 4.4.3. Thus, $F(0) = 0$, implying $u_0 = p_n$ by plugging in $q = 0$. Plugging $c(q) = q^\beta$ back in the constraint, we obtain

$$
p_n + q^\beta = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1}, \tag{D.12}
$$

for every $q \in \operatorname{supp}(F)$. Thus, we have

$$
q = \left(\sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1} - p_n \right)^{1/\beta}. \tag{D.13}
$$

Define

$$
g(F(q), \mathbf{p}) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1} - p_n. \tag{D.14}
$$

Notice that $g(x, \mathbf{p}) = h(x, \mathbf{p}) - p_n$. Then we have $q = g(F(q), \mathbf{p})^{1/\beta}$. From Lemma 4.4.4, we have

$$\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) = \sum_{i=1}^{n} p_i \mathbb{E}\left[\, q_{(i)} \,\right] = \sum_{i=1}^{n} \int_{q \in \mathrm{supp}(F)} (1 - F_{(i)}(x)) dq.$$

Let $f(q)$ be the derivative[7] of $F(q)$ with respect to q.[8]

Thus using Lemma D.3.1, we can rewrite the user welfare function as:[9]

$$\begin{aligned}
\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) &= \int_{q \in \mathrm{supp}(F)} \sum_{i=1}^{n} \left(1 - \sum_{j=1}^{i} \binom{n}{j-1} (1 - F(q))^{j-1} F(q)^{n-j+1}\right) p_i dq \\
&= \int_{q \in \mathrm{supp}(F)} \sum_{i=1}^{n} \left(1 - \sum_{j=1}^{i} \binom{n}{j-1} (1 - F(q))^{j-1} F(q)^{n-j+1}\right) p_i \frac{dq}{dF(q)} dF(q) \\
&= \int_{q \in \mathrm{supp}(F)} \phi(F(q), \mathbf{p}) \frac{1}{f(q)} dF(q) \qquad \text{(Equation (D.8))}
\end{aligned}$$

From the constraints on the condition for being equilibrium in Equation (D.12), we have $p_n + q^\beta = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i} (1 - F(q))^{i-1}$. By differentiating over q, we have

$$\beta q^{\beta - 1} = f(q) \psi(F(q), \mathbf{p}), \tag{D.15}$$

since $\psi(x, \mathbf{p})$ is a derivative of $\left(\sum_{i=1}^{n} p_i \binom{n-1}{i-1} x^{n-i} (1 - x)^{i-1}\right)$ with respect to x due to Lemma D.3.2.

[7] Remark that since F is not guaranteed to be absolute continuous, f does not necessarily coincide with the density function. However, our analysis does not require f to be the density function.

[8] Since F is strictly increasing and atomless, it is differentiable almost everywhere by Lebesgue's differentiation theorem, implying that the derivative of F exists almost everywhere. Thus, since sets of measure zero have no effect on the value of the integral—as the objective function is always finite—it is without loss of generality to assume the differentiability of F. Alternatively, one can restrict the integrals and the corresponding constraints to the set of differentiable points, which would yield the same objective value as the original formulation.

[9] Note that $f(q) \neq 0$ with probability 1 by Proposition 4.4.3.

Finally, plugging $f(q)$ yields

$$\begin{aligned}
\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) &= \int_{q \in \mathrm{supp}(F)} \phi(F(q), \mathbf{p}) \psi(F(q), \mathbf{p}) \frac{1}{\beta q^{\beta-1}} dF(q) \\
&= \frac{1}{\beta} \int_{q \in \mathrm{supp}(F)} \phi(F(q), \mathbf{p}) \psi(F(q), \mathbf{p}) q^{-\beta+1} dF(q) \\
&= \frac{1}{\beta} \int_{q \in \mathrm{supp}(F)} \phi(F(q), \mathbf{p}) \psi(F(q), \mathbf{p}) g(F(q), \mathbf{p})^{-1+1/\beta} dF(q) \\
&= \frac{1}{\beta} \int_0^1 \phi(x, \mathbf{p}) \psi(x, \mathbf{p}) g(x, \mathbf{p})^{1/\beta-1} dx
\end{aligned}$$

Thus, this is entirely a function of $\mathbf{p}$ whose integrand does not explicitly depend on the cumulative distribution function $F(q)$.

Note further that

$$\begin{aligned}
\frac{dg(x, \mathbf{p})^{1/\beta}}{dx} &= \frac{dg(x, \mathbf{p})^{1/\beta}}{dg(x, \mathbf{p})} \frac{dg(x, \mathbf{p})}{dx} \\
&= \frac{1}{\beta} g(x, \mathbf{p})^{1/\beta-1} \frac{dh(x, \mathbf{p})}{dx} \\
&= \frac{1}{\beta} g(x, \mathbf{p})^{1/\beta-1} \psi(x, \mathbf{p}) \qquad \text{(D.16)}
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) &= \int_0^1 \phi(x, \mathbf{p}) \cdot \frac{d}{dx} g(x, \mathbf{p})^{1/\beta} \, dx \\
&= \left[\phi(x, \mathbf{p}) \cdot g(x, \mathbf{p})^{1/\beta} \right]_0^1 - \int_0^1 g(x, \mathbf{p})^{1/\beta} \cdot \frac{d}{dx} \phi(x, \mathbf{p}) \, dx.
\end{aligned}$$

Since we have

$$g(0, \mathbf{p}) = p_n - p_n = 0$$

$$g(1, \mathbf{p}) = p_1 - p_n$$

$$\phi(1, \mathbf{p}) = 0$$

$$\phi(0, \mathbf{p}) = 1$$

we obtain

$$\begin{aligned}
\beta \cdot \mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) &= \left(\beta \phi(x, \mathbf{p}) g(x, \mathbf{p})^{1/\beta} \Big|_0^1 - \beta \int_0^1 \frac{d\phi(x, \mathbf{p})}{dx} g(x, \mathbf{p})^{1/\beta} dx \right) \\
&= \beta \int_0^1 \frac{-d\phi(x, \mathbf{p})}{dx} g(x, \mathbf{p})^{1/\beta} dx \\
&= \beta \int_0^1 \frac{-d\phi(x, \mathbf{p})}{dx} (h(x, \mathbf{p}) - p_n)^{1/\beta} dx.
\end{aligned}$$

Finally by Lemma D.3.4, we obtain

$$\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) = n \int_0^1 h(x, \mathbf{p}) g(x, \mathbf{p})^{1/\beta} dx. \tag{D.17}$$

Now, consider the perturbation on $\mathbf{p}$ that moves small $0 < \delta < p_n$ from p_n to p_1. Writing the resulting vector $\mathbf{p}^{(\delta)}$, for any $x \in [0, 1]$ we have

$$\frac{d}{d\delta} h(x, \mathbf{p}^{(\delta)}) = a_1(x) - a_n(x),$$

and

$$\frac{d}{d\delta}(h(x,\mathbf{p}^{(\delta)})-p_n^{(\delta)})=1+a_1(x)-a_n(x).$$

Thus, we have

$$\begin{aligned}
&\frac{d}{d\delta}\big(h(x,\mathbf{p})(h(x,\mathbf{p})-p_n)^{1/\beta}\big)\\
&=(a_1(x)-a_n(x))(h(x,\mathbf{p})-p_n)^{1/\beta}+\frac{h(x,\mathbf{p})}{\beta}(h(x,\mathbf{p})-p_n)^{1/\beta-1}(1+a_1(x)-a_n(x))\\
&\geq(a_1(x)-a_n(x))(h(x,\mathbf{p})-p_n)^{1/\beta}+\frac{(h(x,\mathbf{p})-p_n)^{1/\beta}}{\beta}(1+a_1(x)-a_n(x)).
\end{aligned}$$

The second term is apparently nonnegative as $a_n(x)\in[0,1]$, and is even strictly positive if $x\in(0,1)$.

Thus, let us focus on the first term. In differentiating D.17 with respect to δ, the first term will appear as an integrand as follows

$$\int_0^1(a_1(x)-a_n(x))(h(x,\mathbf{p})-p_n)^{1/\beta}dx.$$

Notice that $a_1(x)-a_n(x)$ is increasing on x. Since $h(x,\mathbf{p})$ is also increasing on x by Lemma D.5.1, by FKG inequality (Lemma D.3.5), we have

$$\begin{aligned}
\int_0^1(a_1(x)-a_n(x))(h(x,\mathbf{p})-p_n)^{1/\beta}dx&\geq\int_0^1(a_1(x)-a_n(x))dx\int_0^1(h(x,\mathbf{p})-p_n)^{1/\beta}dx\\
&\geq 0.
\end{aligned}$$

Combining, $\mathcal{W}(\boldsymbol{\mu},\mathbf{p})$ only increases by such operation, implying that optimal policy should have $p_n=$

0. Note further that this implies

$$q^{\beta} = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1} = h(F(q), \mathbf{p}), \tag{D.18}$$

Similarly, for the quality, we have

$$\begin{aligned}
\mathcal{Q}(\boldsymbol{\mu}, \mathbf{p}) &= \mathbb{E}_{q\sim\mu_i}[q] \\
&= \int_{q\in\text{supp}(F)} q dF(q) \\
&= \int_{q\in\text{supp}(F)} \left(\sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1} - p_n \right)^{1/\beta} dF(q) \\
&= \int_0^1 g(x, \mathbf{p})^{1/\beta} dx.
\end{aligned}$$

Recalling the definition of $g(x, \mathbf{p}) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1} - p_n$ from Equation (D.14), observe that the coefficient of p_n is negative as p_n's coefficient is $(1-F(q))^{n-1} - 1$, whereas all the other indices have positive coefficient.

Combining, both the user welfare and average quality increases by decreasing p_n, so an optimal policy would have $p_n = 0$.[10]

Therefore, we have $g(x, \mathbf{p}) = h(x, \mathbf{p})$, and thus the average quality and user welfare can finally be written as

$$\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) = n \int_0^1 h(x, \mathbf{p})^{1+1/\beta} dx, \tag{D.19}$$

$$\mathcal{Q}(\boldsymbol{\mu}, \mathbf{p}) = \int_0^1 h(x, \mathbf{p})^{1/\beta} dx. \tag{D.20}$$

[10]There could be boundary regimes where it remains the same by decreasing p_n, *e.g.,* by taking $\alpha = 0$ and $\beta = 1$. In this case, one can easily check that any policy would trivially optimal by Lemma D.3.3.

This concludes that our optimization problem can be written as

$$
\begin{aligned}
(\text{OPT}_2) \quad \underset{\mathbf{p}\in\Delta_m}{\text{maximize}} \quad & \alpha n \int_0^1 h(x,\mathbf{p})^{1+1/\beta} dx + (1-\alpha)\int_0^1 h(x,\mathbf{p})^{1/\beta} dx \\
\text{subject to} \quad & c(q) = \sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1} \quad \forall q \in \text{supp}(F) \\
& p_1 \geq p_2 \geq \ldots \geq p_n, \ \exists i,j : p_i \neq p_j \\
& \sum_{i=1}^{n} p_i = 1 \\
& F \text{ is c.d.f. of symmetric MNE given } \mathbf{p}
\end{aligned}
$$

Since the objective function is purely a function over $\mathbf{p}$, we can simply ignore the first and the last constraints involving F since for any $\mathbf{p}$, we know that there exists an induced symmetric MNE $\bar{\mu}$ and its c.d.f. F that satisfies these constraints. Thus, our optimization becomes

$$
\begin{aligned}
(\text{OPT}) \quad \underset{\mathbf{p}\in\Delta_m}{\text{maximize}} \quad & \alpha n \int_0^1 h(x,\mathbf{p})^{1+1/\beta} + (1-\alpha)\int_0^1 h(x,\mathbf{p})^{1/\beta} \\
\text{subject to} \quad & p_1 \geq p_2 \geq \ldots \geq p_n, \ \exists i,j : p_i \neq p_j \\
& \sum_{i=1}^{n} p_i = 1
\end{aligned}
$$

and it finishes the proof. a

D.3.2 Proof of Proposition 4.5.1

Proof. If $\alpha = 0$, recall that

$$
G(\mathbf{p}) = \int_0^1 h(x,\mathbf{p})^{1/\beta} dx.
$$

We will lower bound $G(\mathbf{p})$ of UNIFORMBUTLAST policy and argue that it can be significantly larger than $G(\text{HARDMAX})$ for sufficiently large n.

Recall from Proposition **??** that

$$G(\text{HARDMAX}) = \alpha\frac{\beta n}{\beta n + n - 1} + (1-\alpha)\frac{\beta}{\beta + n - 1} = \frac{\beta}{\beta + n - 1}.$$

For UNIFORMBUTLAST policy, observe that

$$h(x, \text{UNIFORMBUTLAST}) = \sum_{i=1}^{n} a_i(x)p_i = \sum_{i=1}^{n-1} \frac{a_i(x)}{n-1}.$$

Since $\sum_{i=1}^{n} a_i(x) = 1$ by Binomial theorem, we have

$$h(x, \text{UNIFORMBUTLAST}) = \frac{1}{n-1}(1 - (1-x)^{n-1}).$$

Thus, we have

$$\begin{aligned}
\int_0^1 h(x, \text{UNIFORMBUTLAST})^{1/\beta} dx &= \int_0^1 \left(\frac{1-(1-x)^{n-1}}{n-1}\right)^{1/\beta} dx \\
&= \int_0^1 \frac{(1-(1-x)^{n-1})^{1/\beta}}{(n-1)^{1/\beta}} dx \\
&\geq \int_{1/2}^1 \frac{(1-(1-x)^{n-1})^{1/\beta}}{(n-1)^{1/\beta}} dx \\
&\geq \int_{1/2}^1 \frac{(1-1/2^{n-1})^{1/\beta}}{(n-1)^{1/\beta}} dx \\
&\geq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{(n-1)^{1/\beta}}.
\end{aligned}$$

As $\beta > 1$, we conclude that there exists a sufficiently large n such that

$$G(\textsc{HardMax})/G(\textsc{UniformButLast}) < \varepsilon$$

for any $\varepsilon > 0$. a

D.4 Omitted Proofs in Section 4.6

D.4.1 Proof of Lemma 4.6.5

We require the following property on the Bernstein basis polynomials:

Lemma D.4.1 ([236]). *Define a Bernstein basis polynomial*

$$a_{i,n}(x) = \binom{n}{i} x^{n-i}(1-x)^i \,,$$

and hence $a_i(x) = a_{i-1,n-1}(x)$. *Then, for any nonnegative integer* n *and* $i = 0, \ldots, n$, *it follows that*

$$(a_i(x) - a_{i+1}(x)) = -\frac{1}{n}\frac{da_{i,n}(x)}{dx}.$$

Although the proof can be found in many textbooks [235], we provide the proof for completeness.

Proof. We will compute the left-hand side (LHS) and right-hand side (RHS) individually to argue its equality.

First, we evaluate the LHS:

$$\begin{aligned} a_i(x) - a_{i+1}(x) &= \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1} - \binom{n-1}{i} x^{n-i-1}(1-x)^i \\ &= x^{n-i-1}(1-x)^{i-1}\left[\binom{n-1}{i-1}x - \binom{n-1}{i}(1-x)\right] \end{aligned}$$

Using the identities $\binom{n-1}{i-1} = \frac{i}{n}\binom{n}{i}$ and $\binom{n-1}{i} = \frac{n-i}{n}\binom{n}{i}$, we substitute the coefficients:

$$\begin{aligned} a_i(x) - a_{i+1}(x) &= x^{n-i-1}(1-x)^{i-1}\left[\frac{i}{n}\binom{n}{i}x - \frac{n-i}{n}\binom{n}{i}(1-x)\right] \\ &= \frac{1}{n}\binom{n}{i} x^{n-i-1}(1-x)^{i-1}\left[ix - (n-i)(1-x)\right] \\ &= \frac{1}{n}\binom{n}{i} x^{n-i-1}(1-x)^{i-1}[ix - (n - nx - i + ix)] \\ &= \frac{1}{n}\binom{n}{i} x^{n-i-1}(1-x)^{i-1}[nx - (n-i)]\,. \end{aligned} \tag{D.21}$$

Next, we evaluate the RHS by taking the derivative of $a_{i,n}(x)$:

$$\begin{aligned} -\frac{1}{n}\frac{d}{dx}a_{i,n}(x) &= -\frac{1}{n}\frac{d}{dx}\left[\binom{n}{i}x^{n-i}(1-x)^i\right] \\ &= -\frac{1}{n}\binom{n}{i}\left[(n-i)x^{n-i-1}(1-x)^i - ix^{n-i}(1-x)^{i-1}\right] \\ &= -\frac{1}{n}\binom{n}{i}x^{n-i-1}(1-x)^{i-1}\left[(n-i)(1-x) - ix\right] \\ &= -\frac{1}{n}\binom{n}{i}x^{n-i-1}(1-x)^{i-1}[n - nx - i + ix - ix] \\ &= -\frac{1}{n}\binom{n}{i}x^{n-i-1}(1-x)^{i-1}[n - i - nx] \\ &= \frac{1}{n}\binom{n}{i}x^{n-i-1}(1-x)^{i-1}[nx - (n-i)], \end{aligned}$$

which coincides with Equation (D.21). a

Now we are ready to prove the main theorem.

Proof of Theorem 4.6.5. Consider any probability vector $\mathbf{p}$ in n-dimensional probability simplex, which is possibly not sorted. Recall a generalized notion of $a_i(x)$ as follows: $a_{i,n}(x) := \binom{n}{i} x^{n-i}(1-x)^i$ and note that $a_i(x) = a_{i-1,n-1}(x)$. Then, we will use the following well-known fact from Bernstein basis polynomials [235], presented in Lemma D.4.1:

$$n(a_i(x) - a_{i+1}(x)) = -\frac{da_{i,n}(x)}{dx}. \tag{D.22}$$

Recall that $O(x, \mathbf{p}) = \sum_{i=1}^{n} a_i(x) p_{(i)}$. Let $\sigma(\cdot)$ be the permutation associated with p_i's such that p_i is the $\sigma(i)$-th largest one among $p_1, \ldots, p_n$. Consider two indices i, j such that $p_i \geq p_j$, *i.e.,* $\sigma(i) < \sigma(j)$. Since $p_i \geq p_j$, by Theorem 4.6.4, it suffices to prove that

$$\frac{\partial \int_0^1 O(x, \mathbf{p})^r dx}{\partial p_i} - \frac{\partial \int_0^1 O(x, \mathbf{p})^r dx}{\partial p_j} \geq (\leq)\, 0,$$

depending on the parameter r. Notice that

$$\frac{\partial \int_0^1 O(x, \mathbf{p})^r dx}{\partial p_i} = \int_0^1 r O(x, \mathbf{p})^{r-1} a_{\sigma(i)}(x) dx.$$

Thus, it suffices to prove that for any index i, j with $\sigma(i) < \sigma(j)$:

$$\int_0^1 r O(x, \mathbf{p})^{r-1} (a_{\sigma(i)}(x) - a_{\sigma(j)}(x)) dx \geq (\leq)\, 0.$$

We will prove a stronger statement such that for any index $i = 1, 2, \ldots, n-1$, it follows that

$$\int_0^1 rO(x,\mathbf{p})^{r-1}(a_i(x) - a_{i+1}(x))dx \geq (\leq)\, 0.$$

Note that this immediately finishes the proof. Recalling that $n(a_i(x) - a_{i+1}(x)) = -\frac{da_{i,n}(x)}{dx}$, the left hand side can be written as

$$\int_0^1 r\, O(x,\mathbf{p})^{r-1}(a_i(x) - a_{i+1}(x))\, dx = -\frac{r}{n}\int_0^1 O(x,\mathbf{p})^{r-1}\frac{d}{dx}a_{i,n}(x)\, dx.$$

Now, apply integration by parts:

$$\int_0^1 O(x,\mathbf{p})^{r-1}\frac{d}{dx}a_{i,n}(x)\, dx = \left[O(x,\mathbf{p})^{r-1}a_{i,n}(x)\right]_0^1 - \int_0^1 a_{i+1,n}(x)\frac{d}{dx}O(x,\mathbf{p})^{r-1}\, dx.$$

for $i \in [n-1]$. Note that $a_{i,n}(x)$ is always zero for $i \in [n-1]$ if $x = 0$ or $x = 1$. Further, we have

$$\frac{\partial O(x,\mathbf{p})^{r-1}}{\partial x} = (r-1)O(x,\mathbf{p})^{r-2}\psi(x,\mathbf{p}) \geq 0,$$

by Lemma D.3.2 and $\psi(x,\mathbf{p}) \geq 0$ due to Lemma D.5.1. Thus, $\int_0^1 rO(x,\mathbf{p})^{r-1}(a_i(x) - a_{i+1}(x))dx$ becomes equivalent to

$$\frac{r(r-1)}{n}\int_0^1 O(x,\mathbf{p})^{r-2}\psi(x,\mathbf{p})a_{i,n}(x)dx,$$

which is ≥ 0 if $r \geq 1$ or $r \leq 0$, and otherwise ≤ 0 as all the terms $O(x,\mathbf{p}), \psi(x,\mathbf{p})$ and $a_{i,n}(x)$ are nonnegative. a

D.4.2 Proof of Theorem 4.1.2

Proof. Note first that the OPT in Theorem 4.1.1 enforces $p_n = 0$. To prove the first condition, if $\alpha = 1$, then the objective function becomes $\mathcal{W}(\boldsymbol{\mu}, \mathbf{p})$.

From Equation (D.19), recall that the user welfare is given by

$$\mathcal{W}(\boldsymbol{\mu}, \mathbf{p}) = n \int_0^1 h(x, \mathbf{p})^{1+1/\beta} dx.$$

Since $1 + 1/\beta \geq 1$ for any $\beta > 0$, the integrand is Schur-convex in $\mathbf{p}$ by Theorem 4.6.5. Moreover, since HARDMAX majorizes any $\mathbf{p} \in \Delta_n$ with $p_n = 0$, this implies that HARDMAX maximizes welfare among such strategies.

Similarly, from Equation (D.20) user quality is given by

$$\mathcal{Q}(\boldsymbol{\mu}, \mathbf{p}) = \int_0^1 h(x, \mathbf{p})^{1/\beta} dx.$$

When $\beta \leq 1$, we again have $1/\beta \geq 1$, and so the integrand is Schur-convex in $\mathbf{p}$. Therefore, when $\beta \leq 1$, the overall objective—being a sum of two Schur-convex functions—is also Schur-convex. It follows that HARDMAX is the optimal allocation in this case.

For the remaining conditions, observe that the objective function reduces to $\mathcal{Q}(\boldsymbol{\mu}, \mathbf{p})$ when $\alpha = 0$. If $\beta > 1$, then $1/\beta < 1$, and the integrand becomes Schur-concave in $\mathbf{p}$. In this case, the uniform policy UNIFORMBUTLAST is majorized by every $\mathbf{p} \in \Delta_n$ with $p_n = 0$, and thus UNIFORMBUTLAST is optimal by the definition of Schur-concave functions.

Finally, if $\alpha = 0$ and $\beta = 1$, then the objective function simply reduces to (up to constant)

$$\int_0^1 h(x, \mathbf{p})dx = \int_0^1 a_i(x)p_i dx.$$

Note that

$$\int_0^1 a_i(x)dx = 1/n,$$

by Lemma D.3.3 and hence the objective function is always a constant. Since the sufficient condition to write the optimization problem is to have a nontrivial policy such that $p_i \neq p_j$ for some $i, j \in [n]$ and $p_n = 0$, we finish the proof. a

D.5 Omitted Proofs in Section 4.7

D.5.1 Proof of Lemma 4.7.2

Before proving the main lemma, we require the following property technical property on the monotonicity of $h(x, \mathbf{p})$ with respect to x.

Lemma D.5.1. *$h(x, \boldsymbol{p})$ is strictly increasing in $x \in [0, 1]$ for any nontrivial policy $\boldsymbol{p}$.*

Proof. Observe that $\frac{dh(x,\mathbf{p})}{dx} = \sum_{i=1}^n \frac{da_i(x)}{dx} p_i$. We will use the following Abel's transformation formula [237]:

$$\sum_{k=1}^n f_k g_k = F_n g_n - \sum_{k=1}^{n-1} F_k(g_{k+1} - g_k),$$

where $F_k = \sum_{i=1}^k f_i$ for two real sequences f_k's and g_k's. Plugging in $f_k = da_k(x)/dx$ and $g_k = p_k$,

we obtain

$$\sum_{i=1}^{n} \frac{da_i(x)}{dx} p_i = A'_n(x)p_n - \sum_{k=1}^{n-1} A'_k(x)(p_{k+1} - p_k)$$
$$= A'_n(x)p_n + \sum_{k=1}^{n-1} A'_k(x)(p_k - p_{k+1}),$$

for $A'_k(x) = \sum_{i=1}^{k} a'_i(x)$ where the derivative is with respect to x. Noting that $\sum_{i=1}^{n} a_i(x) = 1$ by binomial theorem, we have $A'_n(x) = 0$.

To compute $A'_k(x)$, recall the generalized notation $a_{i,n}(x)$ defined in Lemma D.4.1:

$$a_{i,n}(x) = \binom{n}{i} x^{n-i}(1-x)^i, \tag{D.23}$$

and the equivalence:

$$a_{i-1,n-1}(x) - a_{i,n-1}(x) = -\frac{1}{n}\frac{da_{i,n}(x)}{dx}.$$

Thus, we have

$$a'_i(x) = a'_{i-1,n-1}(x) = (n-1)(a_{i-1,n-2}(x) - a_{i-1,n-2}(x)).$$

Let us define $a_{-1,k}(x) = 0$ for nonnegative integer k, for ease of exposition. Note, then, that equiva-

lence (D.23) holds for $i = 1$ as well. By telescoping, this implies:

$$\begin{aligned} A'_k(x) &= \sum_{i=1}^{k} a'_i(x) \\ &= \sum_{i=1}^{k} (n-1)(a_{i-1,n-2}(x) - a_{i-1,n-2}(x)) \\ &= (n-1)(a_{k-1,n-2}(x) - a_{-1,n-2}(x) = a_{k-1,n-2}(x). \end{aligned}$$

Thus, we finally have

$$\frac{dh(x, \mathbf{p})}{dx} = \sum_{k=1}^{n-1} a_{k-1,n-2}(x)(p_k - p_{k+1}).$$

Note that this is strictly positive for $x \in (0, 1)$ as we have at least one k such that $p_k > p_{k+1}$ for any nontrivial policy $\mathbf{p}$. This finishes the proof. a

Now the main lemma can be proven in what follows.

Proof of Lemma 4.7.2. If $\beta \leq 1$, then the proof immediately follows from the fact that HARDMAX is optimal from Theorem 4.1.2. Thus, we assume $\beta > 1$.

We will first prove that $K(x, i) = a_i(1-x)$ is STP_{n-1}, where we recall that $a_i(x) = \binom{n-1}{i-1} x^{n-i}(1-x)^{i-1}$.[11] That is, for any $k \in [n-1]$, by Definition 4.7.4, it suffices to prove that

$$\det \begin{pmatrix} a_{i_1}(1-x_1) & \dots & a_{i_1}(1-x_k) \\ \vdots & & \vdots \\ a_{i_k}(1-x_1) & \dots & a_{i_k}(1-x_k) \end{pmatrix} > 0,$$

[11] This is in fact proven by [238], but we provide here the proof for completeness.

for all $1 \leq i_1 < \ldots < i_k \leq n-1$ and $0 \leq x_1 < \ldots < x_k \leq 1$. Denote the matrix inside the determinant by M for simplicity.

Consider $r, s \in [k]$. Note first that

$$a_{i_r}(1-x_s) = \binom{n-1}{i_r-1}(1-x_s)^{n-1}\left(\frac{x_s}{1-x_s}\right)^{i_r-1}.$$

Define $y_s = x_s/(1-x_s)$, then

$$a_{i_r}(1-x_s) = \binom{n-1}{i_r-1}(1-x_s)^{n-1}y_s^{i_r-1}.$$

Factoring out the row constants, we have

$$\det(M) = \left(\prod_{r=1}^{k}\binom{n-1}{i_r-1}\right)\det\begin{pmatrix} (1-x_1)^{n-1}y_1^{i_1-1} & \ldots & (1-x_k)^{n-1}y_k^{i_1-1} \\ \vdots & & \vdots \\ (1-x_1)^{n-1}y_1^{i_k-1} & \ldots & (1-x_k)^{n-1}y_k^{i_k-1} \end{pmatrix}$$

Factoring out the column constants, we have

$$\det(M) = \left(\prod_{r=1}^{k}\binom{n-1}{i_r-1}\right)\left(\prod_{s=1}^{k}(1-x_s)^{n-1}\right)\det\begin{pmatrix} y_1^{i_1-1} & \ldots & y_k^{i_1-1} \\ \vdots & & \vdots \\ y_1^{i_k-1} & \ldots & y_k^{i_k-1} \end{pmatrix}$$

By taking transpose in the remaining matrix and applying Theorem 4.7.9, we conclude that $\det(M) > 0$. Since this holds for arbitrary choice of $x_1, \ldots, x_k$ and $i_1, \ldots, i_k$, this concludes that $K(x, i)$ is STP_{n-1}.

Now we will prove that the diminishing property of totally positive matrix $K(x,i)$ allows us to show the quasi-convexity of the sequence of partial derivatives. Let us define $q(x) = f(x) + g(x)$, where

$$f(x) = n\alpha(1+1/\beta)h(x,\mathbf{p})^{1/\beta}$$

$$g(x) = (1-\alpha)/\beta h(x,\mathbf{p})^{1/\beta-1}.$$

Then, by Equation (4.2), we have

$$\begin{aligned} d_i &= \int_0^1 q(x)a_i(x)dx \\ &= \int_0^1 q(1-y)a_i(1-y)dy. \end{aligned}$$

By Lemma 4.7.3, we know that $q(x)$ is quasiconvex over $x \in [0,1]$. This implies that $q(1-y)$ is also quasiconvex on $y \in [0,1]$ because quasiconvexity on a single dimension is preserved under affine transformations of the input variable.

Thus, by Lemma 4.7.8, $q(1-y)-\lambda$ has at most two sign changes for every $\lambda \in \mathbb{R}$, and whenever it has exactly two sign changes, it must be $(+,-,+)$.

Recalling the definition of d_i and $K(x,i) = a_i(1-x)$:

$$d_i = \int_0^1 q(1-y)a_i(1-y)dy = \int_0^1 K(y,i)q(1-y)dy.$$

Subtracting $\lambda' \in \mathbb{R}$:

$$d_i - \lambda' = -\lambda' + \int_0^1 K(y,i)q(1-y)dy = \int_0^1 K(y,i)(q(1-y) - n\lambda')dy,$$

where the second equality follows from Lemma D.3.3, as we show $\int_0^1 a_i(x)dx = 1/n$ for any $i \in [n]$ and $K(x,i) = a_i(1-x)$.

Taking $\lambda = n\lambda'$, we know that $q(1-y) - \lambda$ has at most two sign changes, and the pattern is exactly $(+,-,+)$ if it has exactly two sign changes. Further, since we know that $K(y,i)$ is STP_{n-1}, we can apply Theorem 4.7.7 as $S^-(q(1-y)) \leq 2$ if $n \geq 3$ and conclude that $S^+(d_i - \lambda') \leq 2$ for any $\lambda' \in \mathbb{R}$. This concludes that $d_i - \lambda'$ always have at most two sign changes assigning zeroes arbitrary signs, and the pattern is $(+,-,+)$ if it has exactly two sign changes.

We will first prove that $d_i = d_{i+1}$ can only happen at the slope-changing index, to be defined shortly, or for $i = n-2$ if d_i's are monotone decreasing without slope-changing index.

Suppose first that there exists an index k such that the slope changes, *i.e.,* monotonicity pattern changes. Let k denote the index at which the monotonicity pattern changes, i.e.,

$$d_1 \geq d_2 \geq \ldots \geq d_k \leq d_{k+1} \leq \ldots \leq d_{n-1}.$$

Suppose there exists an index $i \leq k-2$ or $i \geq k+1$ such that $d_i = d_{i+1}$. Then, by choosing $\lambda' = d_i$, and assigning the signs appropriately,[12] we obtain at least four sign changes in the resulting sequence, implying $S^+(\lambda') \geq 3$, which contradicts our assumptions. On the other hand, suppose $d_1 \geq d_2 \geq \ldots \geq d_{n-1}$, *i.e.,* slope does not change. Similarly, if $d_i = d_{i+1}$ for $1 < i < n-1$, then by

[12]Formally, one can assign $(+,-,+,-)$ for $i \leq k-2$, and $(-,+,-,+)$ for $i \geq k+1$.

choosing $\lambda' = d_i$ and similarly assigning signs for zeroes as above, we obtain $S^+(d_i - \lambda') \geq 3$. If $i = 1$, then by setting $\lambda' = d_1$, sign changes become $(\cdot, \cdot, -)$, which contradicts as the sign pattern should be $(+, -, +)$ in case of exact two changes. Similarly, if the entire sequence is monotone increasing, then $d_i = d_{i+1}$ can only occur if $i = 1$. Therefore, $d_i = d_{i+1}$ can only occur for $i = k - 1$ or $i = k$, or $i = n - 1$ if it is monotone decreasing sequence, or $i = 1$ if it is monotone increasing sequence.

Finally, we will prove that if k is the slope-changing index, then exactly one of $d_{k-1} = d_k$ or $d_k = d_{k+1}$ happens. Suppose there exists a slope-changing index k, and assume $d_{k-1} = d_k = d_{k+1}$. Setting $\lambda' = d_k$, we can assign the signs of the zeros at $d_{k-1} - \lambda'$, $d_k - \lambda'$, and $d_{k+1} - \lambda'$ as $(-, +, -)$, resulting in a sign pattern of the form $(+, -, +, -, +)$, which again leads to a contradiction. Thus, there can be at most two consecutive equal values among (d_{k-1}, d_k) or (d_k, d_{k+1}), and the rest of the sequence must be strictly monotonic on either side. This completes the proof.

a

D.5.2 Proof of Lemma 4.7.3

Proof of Lemma 4.7.3. Recall that $\psi(x, \mathbf{p})$ is a derivative of $h(x, \mathbf{p})$ with respect to x by Lemma D.3.2. Thus, $\psi(x, \mathbf{p}) = dh(x, \mathbf{p})/dx \geq 0$ from Lemma D.5.1. Differentiating $q(x)$ over x, we obtain

$$\begin{aligned}\frac{dq(x)}{dx} &= n\alpha(1 + 1/\beta)(1/\beta)h(x, \mathbf{p})^{1/\beta - 1}\psi(x, \mathbf{p}) + (1 - \alpha)(\frac{1}{\beta} - 1)\frac{1}{\beta}h(x, \mathbf{p})^{1/\beta - 2}\psi(x, \mathbf{p}) \\ &= \psi(x, \mathbf{p})h(x, \mathbf{p})^{1/\beta - 2}(w_1 h(x, \mathbf{p}) + w_2),\end{aligned}$$

where we write $w_1 = n\alpha(1 + 1/\beta)(1/\beta)$ and $w_2 = (1 - \alpha)(\frac{1}{\beta} - 1)\frac{1}{\beta}$.

Note that $w_1 \geq 0$ and $w_2 \leq 0$ as we have $\beta > 1$. Thus $dq(x)/dx$ is nonnegative if and only if $w_1 h(x, \mathbf{p}) + w_2 \geq 0$. Since $h(x, \mathbf{p})$ is increasing function over x by Lemma D.5.1, this concludes

that $q(x)$ is either increasing/decreasing function, or decreasing then increasing function, *i.e.,* it is quasiconvex. a

D.5.3 Proof of Theorem 4.1.3

Proof of Theorem 4.1.3. For a policy $\mathbf{p}$ to be optimal, the Karush-Kuhn-Tucker (KKT) condition should hold as the constraints are linear. We will ignore p_n as it is always zero in the optimum. First, the Lagrangian can be written as:

$$L(\mathbf{p}, \nu, \lambda) = G(\mathbf{p}) + \lambda(1 - \sum_{i=1}^{n-1} p_i) + \sum_{i=1}^{n-2} \nu_i(p_i - p_{i+1}) + \nu_{n-1}p_{n-1},$$

where $\nu_i \geq 0$ for $i \in [n-1]$ and $\lambda \in \mathbb{R}$. Writing down KKT condition, we obtain

$$\begin{aligned}
d_1 = \frac{\partial G(\mathbf{p})}{\partial p_1} &= \lambda - \nu_1 \\
d_i = \frac{\partial G(\mathbf{p})}{\partial p_i} &= \lambda + \nu_{i-1} - \nu_i, \ i = 2, 3, \ldots, n-1 \\
\nu_i(p_i - p_{i+1}) &= 0, \ i = 1, 2, \ldots, n-2 \\
\nu_{n-1}p_{n-1} &= 0 \\
\lambda(p_1 + \ldots + p_n - 1) &= 0 \\
p_1 \geq p_2 \geq \ldots \geq p_{n-1} &\geq 0 \\
\nu_i \geq 0, \ i &= 1, 2, \ldots, n-1
\end{aligned}$$

According to Lemma 4.7.2, the sequence $(d_1, \ldots, d_{n-1})$ can only have either of the following

structures for some $k \in [n-1]$:

1. $d_1 > d_2 > \ldots > d_k \leq d_{k+1} < \ldots < d_{n-1}$

2. $d_1 > d_2 > \ldots > d_{n-2} \geq d_{n-1}$

3. $d_1 \leq d_2 < d_3 < \ldots, d_{n-1}$.

Let k be the slope-changing index in the former two cases.

Overall, given the properties of d_i's above, we will consider two scenarios. First we will show that if $p_1 \neq p_2$, then we cannot have i such that $p_i \neq p_{i+1}$ for $i = 2, \ldots, n-2$, concluding that $p_1 > p_2 = \ldots = p_{n-1}$. Then, we will show that if $p_1 = p_2$, then we always have $p_2 = \ldots = p_{n-1}$.

Scenario 1: $\mathbf{p_1 \neq p_2}$. Assume first that $p_1 \neq p_2$. Then, we have $\nu_1 = 0$. Thus, we have

$$d_1 = \lambda$$

$$d_2 = \lambda - \nu_2.$$

We will deal with two cases where $p_2 = p_3$ or $p_i = p_{i+1}$ for $i = 3, \ldots, n-2$ separately.

Scenario 1a: $\mathbf{p_2 \neq p_3}$. Suppose, for the sake of contradiction, that $p_2 \neq p_3$. Then, we must have $\nu_2 = 0$, which implies

$$d_1 = \lambda,$$

$$d_2 = \lambda,$$

$$d_3 = \lambda - \nu_3.$$

This implies that the slope-changing index is $k = 1$. However, this leads to a contradiction: we have

$d_3 = \lambda - \nu_3 \leq \lambda = d_2$ since $\nu_3 \geq 0$, but since $k = 1$ is the slope-changing index, it should have been $d_2 < d_3$. Therefore, we conclude that if $p_1 \neq p_2$, then it must be that $p_2 = p_3$.

Scenario 1b: $\mathbf{p_i \neq p_{i+1}}$ for $\mathbf{i \geq 3}$. Now, again for the sake of contradiction, suppose that the first index i such that $p_i \neq p_{i+1}$ is greater than or equal to 3. That is,

$$p_1 \neq p_2 = p_3 = \ldots = p_i \neq p_{i+1},$$

for some $i \geq 3$. Then $\nu_i = 0$ and we have the following series of equations:

$$\begin{aligned} d_1 &= \lambda \\ d_2 &= \lambda - \nu_2 \\ d_3 &= \lambda + \nu_2 - \nu_3 \\ &\vdots \\ d_i &= \lambda + \nu_{i-1} \\ d_{i+1} &= \lambda - \nu_{i+1}. \end{aligned}$$

If $\nu_{i+1} \neq 0$ and $\nu_{i-1} \neq 0$, then it is a contradiction as the d sequence increases then decreases. Thus, either of ν_{i-1} or ν_{i+1} is zero.

If $\nu_{i+1} = 0$, then we should again have $\nu_{i-1} = 0$ as otherwise the d sequence increases from $d_1 = \lambda$ to $d_i = \lambda + \nu_{i-1}$, then decreases to $d_{i+1} = \lambda - \nu_{i+1}$. Then, notice that by setting $\nu_{i+1} = \nu_{i-1} = 0$,

we have

$$
\begin{aligned}
d_1 &= \lambda \\
d_2 &= \lambda - \nu_2 \\
d_3 &= \lambda + \nu_2 - \nu_3 \\
&\vdots \\
d_{i-1} &= \lambda + \nu_{i-2} \\
d_i &= \lambda \\
d_{i+1} &= \lambda,
\end{aligned}
$$

which is a contradiction as i should be the slope-changing index and thus it should be the case that $\lambda = d_1 > d_2 > \ldots > d_i = d_{i+1} = \lambda$, but we have $d_1 = \lambda \leq d_{i-1} = \lambda + \nu_{i-2} \geq d_i = d_{i+1}$.

Therefore, we have $\nu_{i-1} = 0$. In this case, again, we have

$$
\begin{aligned}
d_1 &= \lambda \\
d_2 &= \lambda - \nu_2 \\
d_3 &= \lambda + \nu_2 - \nu_3 \\
&\vdots \\
d_i &= \lambda \\
d_{i+1} &= \lambda - \nu_{i+1},
\end{aligned}
$$

and thus it should be the case that $\lambda = d_1 > d_2 > \ldots \geq d_k < \ldots < d_i = \lambda < d_{i+1}$ since slope-

changing needs to occur strictly before the index i, which is a contradiction as $\lambda = d_i \geq d_{i+1} = \lambda - \nu_{i+1}$. Thus, if $p_1 \neq p_2$, we always have $p_1 \neq p_2 = \ldots = p_{n-1} \geq p_n = 0$.

Scenario 2: $\mathbf{p_1 = p_2}$. Now the remaining case is when $p_1 = p_2$. Similarly, we will separately deal with two cases: (i) when $p_i \neq p_{i+1}$ for some $i = 3, \ldots, n-2$, and (ii) when $p_2 \neq p_3$. Let $i \geq 2$ be the first index such that $p_i \neq p_{i+1}$, *i.e.,* $\nu_i = 0$, for $i \in [n-2]$.

Scenario 2a: $\mathbf{p_1 = p_2 > \ldots > p_i = p_{i+1}}$ for $i = 3, \ldots, n-2$. In this case, we have

$$
\begin{aligned}
d_1 &= \lambda - \nu_1 \\
d_2 &= \lambda + \nu_1 - \nu_2 \\
d_3 &= \lambda + \nu_2 - \nu_3 \\
&\vdots \\
d_{i-1} &= \lambda + \nu_{i-2} - \nu_{i-1} \\
d_i &= \lambda + \nu_{i-1} \\
d_{i+1} &= \lambda - \nu_{i+1}.
\end{aligned}
$$

For these equations to be true under the condition on d_i's, it should be the case that either of ν_{i-1} and ν_{i+1} is zero as otherwise it increases from d_1 then decreases to d_{i+1}. If $\nu_{i-1} = 0$, we have $d_{i-1} = \lambda + \nu_{i-2}$, which is again a contradiction as we have $d_1 \leq d_{i-1} \geq d_i \geq d_{i+1}$. If $\nu_{i+1} = 0$, then we should have $d_i = d_{i+1}$ as otherwise it increases from $d_1 = \lambda - \nu_1$ then decreases to λ. Thus, $d_i = d_{i+1} = \lambda$. Again, in this case, we have $d_{i-1} = \lambda + \nu_{i-2}$, which is a contradiction.

Scenario 2b: $\mathbf{p_1 = p_2 > p_3}$. In this case, $\nu_2 = 0$ and we have

$$d_1 = \lambda - \nu_1$$

$$d_2 = \lambda + \nu_1$$

$$d_3 = \lambda - \nu_3$$

$$d_4 = \lambda + \nu_3 - \nu_4$$

Thus, it should be the case that $\nu_1 = 0$ as otherwise $d_1 < d_2 \geq d_3$. Then, the resulting equations are:

$$d_1 = \lambda$$

$$d_2 = \lambda$$

$$d_3 = \lambda - \nu_3$$

$$d_4 = \lambda + \nu_3 - \nu_4,$$

and thus the slope-changing index should be $k = 1$. Thus, after $k + 1 = 2$, it should always strictly increase, which is a contradiction to $d_3 = \lambda - \nu_3$. Therefore, if $p_1 = p_2$, we always have $p_1 = p_2 = \ldots = p_{n-1}$.

Combining, the only possible cases are:

$$p_1 > p_2 = \ldots = p_{n-1} \geq p_n = 0$$

$$p_1 = p_2 = \ldots = p_{n-1} \geq p_n = 0,$$

and it finishes the proof.

a

D.6 Omitted Proofs in Section 4.8

Before presenting the proofs, we need the following technical lemma, the proof of which is omitted as it is an elementary algebra.

Lemma D.6.1. *For $r \geq 1$ and $a, b \in [0,1]$, $|a^r - b^r| \leq r\,|a-b|$. For $r \in (0,1]$ and $a, b \geq 0$, $|a^r - b^r| \leq |a-b|^r$.*

D.6.1 Proof of Lemma 4.8.1

Proof of Lemma 4.8.1. For each fixed x, $p_1 \mapsto h(x;p_1) = c_0(x) + c_1(x)p_1$ is affine and $t \mapsto t^r$ is increasing for $r > 0$; hence $\sup_{p_1 \in [\ell,u]} h(x;p_1)^r = \max\{h(x;\ell)^r, h(x;u)^r\}$. Integrating in x, and summing the two exponents $r \in \{1 + 1/\beta,\ 1/\beta\}$ with weights αn and $1-\alpha$, gives the upper bound. The lower bound follows since $L(I)$ is the objective at boundary points. a

D.6.2 Proof of Lemma 4.8.2

Proof of Lemma 4.8.2. Let $p^\diamond \in \{\ell, u\}$ be such that $L(I) = G(p^\diamond)$ and let p° be the other endpoint. By the definition of $U(I)$ and $L(I)$, linearity of the integral, and the inequality $\max\{A,B\} - A \leq |B-A|$

(valid for all real A, B), we have

$$\begin{aligned} U(I) - L(I) &= \alpha n \int_0^1 \Big(\max\{h(x;p^\circ)^{1+1/\beta}, h(x;p^\diamond)^{1+1/\beta}\} - h(x;p^\diamond)^{1+1/\beta} \Big)\, dx \\ &\quad + (1-\alpha) \int_0^1 \Big(\max\{h(x;p^\circ)^{1/\beta}, h(x;p^\diamond)^{1/\beta}\} - h(x;p^\diamond)^{1/\beta} \Big)\, dx \\ &\le \alpha n \int_0^1 \big| h(x;p^\circ)^{1+1/\beta} - h(x;p^\diamond)^{1+1/\beta} \big|\, dx \\ &\quad + (1-\alpha) \int_0^1 \big| h(x;p^\circ)^{1/\beta} - h(x;p^\diamond)^{1/\beta} \big|\, dx. \end{aligned}$$

Define $\Delta_r(x) := \big| h(x;p^\circ)^r - h(x;p^\diamond)^r \big|$ for $r > 0$. Then

$$U(I) - L(I) \;\le\; \alpha n \int_0^1 \Delta_{1+1/\beta}(x)\, dx + (1-\alpha) \int_0^1 \Delta_{1/\beta}(x)\, dx. \tag{D.24}$$

Because $h(x;p_1) \in [0,1]$ and $p_1 \mapsto h(x;p_1)$ is affine, the increment across the interval $I = [\ell, u]$ is

$$h(x;u) - h(x;\ell) = c_1(x)\,(u-\ell) = c_1(x)\,|I|.$$

For $r \ge 1$ and $a, b \in [0,1]$, $|a^r - b^r| \le r\,|a-b|$ and for $r \in (0,1]$ and $a, b \ge 0$, $|a^r - b^r| \le |a-b|^r$ by Lemma D.6.1. Applying these with $(a,b) = (h(x;u), h(x;\ell))$ yields

$$\Delta_{1+1/\beta}(x) \le \Big(1 + \tfrac{1}{\beta}\Big)\, |c_1(x)|\, |I|, \qquad \Delta_{1/\beta}(x) \le |c_1(x)|^{1/\beta}\, |I|^{1/\beta}.$$

Substituting the bounds for $\Delta_{1+1/\beta}$ and $\Delta_{1/\beta}$ into (D.24),

$$U(I) - L(I) \;\le\; \alpha n \Big(1 + \tfrac{1}{\beta}\Big) \int_0^1 |c_1(x)|\, dx\; |I| \;+\; (1-\alpha) \int_0^1 |c_1(x)|^{1/\beta}\, dx\; |I|^{1/\beta},$$

which is the claimed inequality with C_1 and C_2 as stated. a

D.6.3 Proof of Lemma 4.8.4

Proof of Lemma 4.8.4. (i) $L^\star$ is the maximum of objective values at feasible endpoints, hence a valid global lower bound. (ii) Let $I^\star$ be any interval (among the descendants of I_0) that contains a global maximizer; Lemma 4.8.1 gives $U(I^\star) \geq \sup_{p \in I^\star} G(p) \geq \mathrm{OPT}$, and $I^\star$ remains in $\mathcal{A}$ until pruned or split, so $\max_{I \in \mathcal{A}} U(I) \geq U(I^\star) \geq \mathrm{OPT}$. a

D.6.4 Proof of Theorem 4.8.6

Proof of Theorem 4.8.6. Let $I^\star$ be any active interval containing a maximizer. If $|I^\star| \leq \delta(\varepsilon)$, then by Lemma 4.8.2, $U(I^\star) - L(I^\star) \leq \varepsilon$. Since $\mathrm{OPT} \leq U(I^\star)$ and $L^\star \geq L(I^\star)$, we get $\mathrm{OPT} - L^\star \leq \varepsilon$. The loop exits only when $\max_{I \in \mathcal{A}} U(I) \leq L^\star + \varepsilon$; combined with Proposition 4.8.4,

$$\mathrm{OPT} \leq \max_I U(I) \leq L^\star + \varepsilon \quad \Rightarrow \quad G(p^\star) = L^\star \geq \mathrm{OPT} - \varepsilon.$$

If $|I^\star| > \delta(\varepsilon)$, it cannot be pruned because Proposition 4.8.4 gives $U(I^\star) \geq \mathrm{OPT} > L^\star + \varepsilon$. Hence it must be split. After k bisections of the interval containing the maximizer, the length is $\leq D\,2^{-k}$. The stated bound on k ensures $D\,2^{-k} \leq \delta(\varepsilon)$, completing the proof. To see the time complexity, note that a uniform partition of $\mathcal{D}$ with mesh $\delta(\varepsilon)$ has $O(D/\delta(\varepsilon))$ intervals, hence the total number of nodes explored by the algorithm is

$$O\Big(\tfrac{C_1}{\varepsilon}\Big) \; + \; O\Big(\tfrac{C_2}{\varepsilon}\Big)^{\beta},$$

dominated by the second term when $\beta > 1$, which is polynomial in $1/\varepsilon$ given constant β. a

D.6.5 Proof of Lemma 4.8.7

Proof of Lemma 4.8.7. For a monotone increasing $f : [0,1] \to \mathbb{R}$, the upper–minus–lower Riemann sum on the uniform grid $\{j/m\}_{j=0}^{m}$ equals $\frac{1}{m}\sum_{j=1}^{m}[f(j/m) - f((j-1)/m)] = \frac{f(1)-f(0)}{m}$. The integral lies between the two sums, proving the bound. Monotonicity of f_r follows from Lemma D.5.1 and $r > 0$. a

D.7 Omitted Proofs in Section 4.9

D.7.1 Proof of Theorem 4.1.5

Proof of Theorem 4.1.5. We first prove the general statement for arbitrary polynomial with nonnegative cofficients. It suffices to prove that the resulting sequence of partial derivatives with respect to p_i's has quasiconvexity by Lemma 4.7.2 under the new objective function $\mathcal{G}(\bar{\mu}, \mathbf{p})$.

Since we deal with symmetric MNE $\bar{\mu}$, we will simply drop the index i in q_i and μ_i. Using (D.13),[13] for any $j \in [n]$, we have

$$\begin{aligned}
\mathbb{E}_{q\sim\mu}\left[e_j q^{k_j}\right] &= \int_{q\in\mathrm{supp}(F)} e_j q^{k_j} dF(q) \\
&= \int_{q\in\mathrm{supp}(F)} e_j \left(\left(\sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1}\right)^{k_j}\right)^{1/\beta} dF(q) \\
&= \int_{q\in\mathrm{supp}(F)} e_j \left(\sum_{i=1}^{n} p_i \binom{n-1}{i-1} F(q)^{n-i}(1-F(q))^{i-1}\right)^{k_j/\beta} dF(q) \\
&= \int_0^1 e_j h(x,\mathbf{p})^{k_j/\beta} dx.
\end{aligned}$$

[13] Note that one can analogously prove $p_n = 0$, and we omit the detail due to its redundancy.

Note that the partial derivative d_i of the objective function with respect to p_i can be written as:

$$\begin{aligned} d_i = \frac{\partial \mathcal{G}(x, \mathbf{p})}{\partial p_i} &= \frac{\partial}{\partial p_i} \sum_{j=1}^{m} \int_0^1 e_j h(x, \mathbf{p})^{k_j/\beta} dx \\ &= \sum_{j=1}^{m} \int_0^1 \frac{j}{\beta} e_j h(x, \mathbf{p})^{k_j/\beta - 1} a_i(x) dx \\ &= \int_0^1 \left(\sum_{j=1}^{m} \frac{j}{\beta} e_j h(x, \mathbf{p})^{k_j/\beta - 1} \right) a_i(x) dx. \end{aligned}$$

Note that the same term $\sum_{j=1}^{m} \frac{k_j}{\beta} e_j h(x, \mathbf{p})^{k_j/\beta - 1}$ remains the same for every $i \in [n]$.

Given $\mathbf{p}$, define $q(x) = \sum_{j=1}^{m} \frac{k_j}{\beta} e_j h(x, \mathbf{p})^{k_j/\beta - 1}$. We will prove that $q(x)$ is quasiconvex, *i.e.,* it either decreases, increases, or decreases then increases. Note that this concludes the proof.

Let us compute the derivative $q'(x)$:

$$\begin{aligned} q'(x) &= \frac{d}{dx} \left(\sum_{j=1}^{m} e_j \frac{k_j}{\beta} h(x, \mathbf{p})^{\frac{k_j}{\beta} - 1} \right) \\ &= \sum_{j=1}^{m} e_j \frac{k_j}{\beta} \left(\frac{k_j}{\beta} - 1 \right) h(x, \mathbf{p})^{\frac{k_j}{\beta} - 2} \cdot \frac{dh(x, \mathbf{p})}{dx}. \end{aligned}$$

We can factor out the common terms and write:

$$q'(x) = \frac{dh(x, \mathbf{p})}{dx} h(x, \mathbf{p})^{-2} \left(\sum_{j=1}^{m} e_j \frac{k_j}{\beta} \left(\frac{k_j}{\beta} - 1 \right) h(x, \mathbf{p})^{\frac{k_j}{\beta}} \right).$$

By Lemma D.5.1, we know that $h'(x, \mathbf{p}) \geq 0$ for any $x \in [0, 1]$. Further, $h(x, \mathbf{p}) \geq 0$ for any $x \in (0, 1)$. Thus, it suffices to prove that the following quantity

$$\sum_{j=1}^{m} e_j \frac{k_j}{\beta} \left(\frac{k_j}{\beta} - 1 \right) h(x, \mathbf{p})^{\frac{k_j}{\beta}}, \tag{D.25}$$

changes its sign at most once, from $-$ to $+$, as this implies that the original function $q(x)$ is decreases then increases, or just monotone.

Define $c_j = e_j \frac{k_j}{\beta}(\frac{k_j}{\beta} - 1)$. Noting that $h(x, \mathbf{p})$ is strictly increasing and nonnegative by Lemma D.5.1, we will simply write $z = h(x, \mathbf{p})^{1/\beta}$. Then, Equation (D.25) simply becomes a posynomial over z:

$$P(z) = \sum_{j=1}^{m} c_j z^{k_j}.$$

By Descartes' rule of signs, the number of positive real roots of $P(z)$ is less than or equal to the number of sign changes in the sequence of its coefficients $(c_1, \ldots, c_m)$.[14]

Given that $e_j \geq 0$ for all $j \in [m]$, the sign of c_j is entirely determined by the $e_j \frac{k_j}{\beta}(\frac{k_j}{\beta} - 1)$. Note that this changes from $-$ to $+$ based on our assumption. In particular, when e_j's are all nonnegative, this is positive or negative depending on whether $k_j/\beta - 1 > 0$ or < 0, *i.e.,* whether $k_j > \beta$ or $k_j < \beta$. Further, if $c_j \geq 0$, this implies that $j \geq \beta$ and thus $c_k \geq 0$ for any $k \geq j$. Combining, we conclude that the sequence of coefficients $(c_1, \ldots, c_m)$ changes its sign from $-$ to $+$. Thus, the posynomial $P(z)$ has at most one positive real root by Descartes' rule of signs. Further, if we take limit $x \to 0+$, we have $z = h(x, \mathbf{p})^{1/\beta} \to p_n = 0$. In this case, the term with the smallest power of z will dominate in $P(z)$. Thus, the sign changes of $P(z)$ can only happen from $-$ to $+$. This concludes that $q(x)$ is quasiconvex.

a

[14] We note that Descartes' rule is proven to be true with real exponents by [239].

D.7.2 Proof of Corollary 4.9.1

Proof of Corollary 4.9.1. Recall that under the symmetric MNE $\bar{\mu}$, we have

$$\begin{aligned}
\mathcal{W}(\bar{\mu}, \mathbf{p}) &= n \int_0^1 h(x, \mathbf{p})^{1/\beta+1} dx \\
\mathcal{V}(\bar{\mu}, \mathbf{p}) &= \sum_{j=1}^{m} \int_0^1 h(x, \mathbf{p})^{j/\beta} dx \\
\mathcal{U}(\mu, \bar{\mu}_{-i}, \mathbf{p}) &= \frac{1}{n} - \mathbb{E}_{q \in \mu} \left[q^{\beta} \right] \\
&= \frac{1}{n} - n \int_0^1 h(x, \mathbf{p}) dx,
\end{aligned}$$

where the last equation follows from (D.13). Notice, however, that $\int_0^1 h(x, \mathbf{p}) dx = \frac{\sum_{i=1}^{n} p_i}{n} = \frac{1}{n}$, by Lemma D.3.3. Thus, it remains to show that

$$\int_0^1 h(x, \mathbf{p})^{1/\beta+1} dx + \sum_{j=1}^{m} \int_0^1 h(x, \mathbf{p}) h(x)^{j/\beta} dx$$

satisfies the condition stated in Theorem 4.1.5, which is straightforward since all the coefficients are nonnegative. This concludes the proof. a

Proof of Corollary 4.9.2. Let $F_{(1)}$ be the c.d.f. of the largest order statistics among q_i's at symmetric

MNE $\bar{\mu}$ with c.d.f. F. Then, we have

$$\begin{aligned}
\mathbb{E}_{q_i \sim \mu_i, i \in [n]} \left[q_{(1)} \right] &= \int_{q \in \text{supp}(F)} (1 - F_{(1)(q)}) dF(q) \\
&= \int_{q \in \text{supp}(F)} (1 - F(q)^n) dF(q) \\
&= \int_0^1 (1 - x^n) \psi(x, \mathbf{p}) h(x, \mathbf{p})^{1/\beta - 1} dx \\
&= n \int_0^1 x^{n-1} h(x, \mathbf{p})^{1/\beta} dx,
\end{aligned}$$

where the derivations are analogous to the proof of Theorem 4.1.1.

Taking a partial derivative with respect to p_i, we obtain

$$n \int_0^1 x^{n-1} h(x, \mathbf{p})^{1/\beta} a_i(x) dx.$$

Thus, we take $q(x) = n x^{n-1} h(x, \mathbf{p})^{1/\beta}$ to apply Lemma 4.7.2. Then, notice that

$$\begin{aligned}
\frac{dq(x)}{dx} &= n(n-1) x^{n-2} h(x, \mathbf{p})^{1/\beta} + \frac{n}{\beta} x^{n-1} h(x, \mathbf{p})^{1/\beta - 1} \psi(x, \mathbf{p}) \\
&= n x^{n-2} h(x, \mathbf{p})^{1/\beta - 1} ((n-1) h(x, \mathbf{p}) + x/\beta \psi(x, \mathbf{p})).
\end{aligned}$$

Since $\psi(x, \mathbf{p}) \geq 0, h(x, \mathbf{p}) \geq 0$, this is always nonnegative, and thus quasiconvex. Thus, we can apply Lemma 4.7.2 and finish the proof. a

D.7.3 Proof of Corollary 4.9.3

Proof of Corollary 4.9.3. Let us define

$$\Phi(\mathbf{p}, \mu) \;:=\; \mathbb{E}_{q\sim\mu}\left[e^{\lambda q}\right],$$

for the symmetric MNE strategy μ given the policy $\mathbf{p}$. For $M \in \mathbb{N}$, we define the truncated posynomial

$$\Phi_M(\mathbf{p}, \mu) \;:=\; \mathbb{E}\left[\sum_{j=0}^{M} \frac{\lambda^j}{j!} q^j\right] = \sum_{j=0}^{M} \frac{\lambda^j}{j!} \mathbb{E}\left[\, q^j \,\right].$$

For each M, Φ_M is a finite linear combination of moments $\mathbb{E}\left[\, q^j \,\right]$ with nonnegative coefficients $e_j := \lambda^j/j!$. Ordering the exponents as $k_j = j$, the sign pattern $e_j(k_j - \beta)$ exhibits at most one sign change as j increases (it is negative for $j < \beta$, nonnegative for $j \geq \beta$). Hence Theorem 4.1.5 applies and yields a maximizer $\mathbf{p}^{(M)}$ which is piecewise constant with at most two value changes at p_1 or p_n.

Since $q \in [0, 1]$, the Lagrange remainder of the exponential series in its Taylor series gives, uniformly over $q \in [0, 1]$,

$$\left|e^{\lambda q} - \sum_{j=0}^{M} \frac{\lambda^j}{j!}\, q^j\right| \;\leq\; \frac{e^{\lambda}\,(\lambda q)^{M+1}}{(M+1)!} \;\leq\; \frac{e^{\lambda}\,\lambda^{M+1}}{(M+1)!}.$$

Thus, we obtain

$$\sup_{\mathbf{p}\in\Delta} \left|\Phi(\mathbf{p}, \mu) - \Phi_M(\mathbf{p}, \mu)\right| \leq \sup_{q\in[0,1]} \left|e^{\lambda q} - \sum_{j=0}^{M} \frac{\lambda^j}{j!}\, q^j\right| \leq \frac{e^{\lambda}\lambda^{M+1}}{(M+1)!}.$$

The sequence $\{\mathbf{p}^{(M)}\}$ lies in Δ, which is a compact feasible set of policies, so it has a convergent

subsequence $\mathbf{p}^{(M_k)} \to \mathbf{p}^\star$. Since we know that $\Phi_M \to \Phi$ uniformly on Δ, so

$$\Phi\big(\mathbf{p}^{(M_k)}\big) \;\geq\; \sup_{\mathbf{p}\in\Delta} \Phi_{M_k}(\mathbf{p}) - \frac{e^\lambda \lambda^{M_k+1}}{(M_k+1)!} \;\geq\; \Phi_{M_k}(\mathbf{p}) - \frac{e^\lambda \lambda^{M_k+1}}{(M_k+1)!}\,,$$

Hence, letting $k \to \infty$ yields $\Phi(\mathbf{p}^\star) \geq \Phi(\mathbf{p})$ for all $\mathbf{p} \in \Delta$, *i.e.,* $\mathbf{p}^\star$ is optimal for Φ. Since $\mathbf{p}^\star$ is a convergent sequence of $\mathbf{p}^{(M)}$ with the structure $p_1 \geq p_2 = \ldots = p_{n-1} \geq p_n = 0$, the resulting sequence $\mathbf{p}^\star$ also inherits the same structure, as desired.

Finally, it is immediate to extend the analysis to any convex combination of exponential utilities with different λ's, since the uniform convergence analogously carries over and the structural result also follows as the sign of $e_j(k_j - \beta)$ only depends on whether $k_j \geq \beta$ or not. a

Appendix E: Gains-from-trade in Bilateral Trade with a Broker

E.1 Proof of Theorem 6.1.8

We first prove the lower bounds and then present the worst-case instances that match the lower bounds.

E.1.1 Lower bounds

By Lemma 6.4.1, we can assume that $[a_1, b_1] = [0, 1]$ without loss of generality. Now, let F and G be the uniform distributions on $[0, 1]$ and $[a, b]$, respectively, *i.e.,* let

$$F(v) = \begin{cases} v & \text{if } v \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

and

$$G(c) = \begin{cases} \dfrac{c - a}{b - a} & \text{if } c \in [a, b] \\ 0 & \text{otherwise} \end{cases}.$$

(Note that $[a, b]$ might not lie completely in $[0, \infty)$, since we normalize F and G so that F is supported on $[0, 1]$.) It is straightforward to check that both F and G are regular. Thus, by Theorem 6.1.7, the gains-from-trade of any BNIC and IR mechanism that maximizes the broker's expected profit is given

	Subcase				
Case	(a)	(b)	(c)	(d)	(e)
1	1	2/3	2/3	2/3	
2			4/7	4/7	1/2
3			4/7		1/2
4				4/7	1/2
5					1/2

Table E.1: Summary of Lower Bounds

by

$$\begin{aligned}
\text{GFT} &= \iint (v-c)\mathbf{1}\left\{v - c \geq \frac{1-F(v)}{F'(v)} + \frac{G(c)}{G'(c)}\right\} dF(v)\, dG(c) \\
&= \iint (v-c)\mathbf{1}\left\{v - c \geq (1-v) + (c-a)\right\} dF(v)\, dG(c) \\
&= \iint (v-c)\mathbf{1}\left\{v - c \geq T_a\right\} dF(v)\, dG(c),
\end{aligned}$$

where $T_a = \dfrac{1-a}{2}$.

We will now complete the proof by considering the following cases based on the values of a and b:

Case 1: $b \leq 0$

Case 2: $a \leq 0 < b \leq 1$

Case 3: $a \leq 0 < 1 < b$

Case 4: $0 < a < b \leq 1$

Case 5: $0 < a < 1 < b$

Case 6: $a \geq 1$

We will further break down each case (except for Case 6) into the following subcases based on the values of $a + T_a$ and $b + T_a$:

Subcase (a): $b + T_a \leq 0$

Subcase (b): $a + T_a \leq 0 < b + T_a \leq 1$

Subcase (c): $a + T_a \leq 0 < 1 < b + T_a$

Subcase (d): $0 < a + T_a < b + T_a \leq 1$

Subcase (e): $0 < a + T_a < 1 < b + T_a$

Note that not every case-subcase combination is possible, as we will see when we investigate each case-subcase combination below. Table E.1 summarizes the lower-bounds for each case-subcase combination. (The empty cells correspond to impossible case-subcase combinations.)

We will now derive the lower bounds in Table E.1. The following two tables will be helpful for our analysis. Specifically, Table E.2 summarizes FB-GFT for each case.

Case	Case Description	FB-GFT	Derivation
1	$b \leq 0$	$\frac{1-a-b}{2}$	1
2	$a \leq 0 < b \leq 1$	$\frac{3a^2-3a+b^3-3b^2+3b}{6(b-a)}$	2
3	$a \leq 0 < 1 < b$	$\frac{3a^2-3a+1}{6(b-a)}$	3
4	$0 < a < b \leq 1$	$\frac{a^2+ab+b^2-3a-3b+3}{6}$	4
5	$0 < a < 1 < b$	$\frac{(1-a)^3}{6(b-a)}$	5

Table E.2: First-Best Gains-from-Trade by Case

Furthermore, Table E.3 summarizes the GFT of the broker's optimal mechanism in each subcase.

Subcase	Subcase Description	GFT	Derivation
(a)	$b + T_a \leq 0$	$\dfrac{1-a-b}{2}$	1
(b)	$a + T_a \leq 0 < b + T_a \leq 1$	$\dfrac{(a+b-1)(a^2-4ab+10a+4b^2-8b+1)}{24(b-a)}$	6
(c)	$a + T_a \leq 0 < 1 < b + T_a$	$\dfrac{(3a-1)^2}{24(b-a)}$	7
(d)	$0 < a + T_a < b + T_a \leq 1$	$\dfrac{(a+2b-3)^2}{24}$	8
(e)	$0 < a + T_a < 1 < b + T_a$	$\dfrac{(1-a)^3}{12(b-a)}$	9

Table E.3: Gains-from-Trade of Broker's Optimal Mechanism by Subcase

We now prove the desired lower bounds in what follows.

Case 1: $b \leq 0$**.**

Case 1(a): $b + T_a \leq 0$

We have that

$$\frac{\text{GFT}}{\text{FB-GFT}} = 1.$$

Case 1(b): $a + T_a \leq 0 < b + T_a \leq 1$

We seek to lower bound

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{-a^2 + 4ab - 10a - 4b^2 + 8b - 1}{12(b-a)}.$$

Taking the derivative with respect to b, we obtain

$$\begin{aligned}
&\frac{\partial}{\partial b}\left[\frac{-a^2 + 4ab - 10a - 4b^2 + 8b - 1}{12(b-a)}\right] \\
&= \frac{12(b-a)(4a - 8b + 8) - 12(-a^2 + 4ab - 10a - 4b^2 + 8b - 1)}{(12(b-a))^2} \\
&= \frac{(4ab - 8b^2 + 8b - 4a^2 + 8ab - 8a) + (a^2 - 4ab + 10a + 4b^2 - 8b + 1)}{12(b-a)^2} \\
&= \frac{8ab - 4b^2 - 3a^2 + 2a + 1}{12(b-a)^2} \\
&= \frac{(-a + 2b + 1)(3a - 2b + 1)}{12(b-a)^2}.
\end{aligned}$$

Now, since $-a + 2b + 1 \geq 0$ and $3a - 2b + 1 \leq 0$ whenever $a \in (-\infty, -1]$ and $b \in \left(\frac{a-1}{2}, \frac{a+1}{2}\right]$, $\frac{\text{GFT}}{\text{FB-GFT}}$ is an decreasing function with respect to b on $a \in (-\infty, -1]$ and $b \in \left(\frac{a-1}{2}, \frac{a+1}{2}\right]$.

Thus,

$$\begin{aligned}\frac{\text{GFT}}{\text{FB-GFT}} &= \frac{-a^2+4ab-10a-4b^2+8b-1}{12(b-a)} \\ &\geq \frac{-a^2+(2a^2+2a)-10a-(a^2+2a+1)+(4a+4)-1}{6-6a} \\ &= \frac{2-6a}{6-6a} \\ &\geq \frac{2-6(-1)}{6-6(-1)} \\ &= \frac{2}{3},\end{aligned}$$

since $\frac{2-6a}{6-6a}$ is decreasing on $(-\infty,-1]$.

Case 1(c): $a+T_a \leq 0 < 1 < b+T_a$

Since $(b-a)(1-a-b)$ is increasing with respect to b on $(-\infty,0]$, we have that

$$\begin{aligned}\frac{\text{GFT}}{\text{FB-GFT}} &= \frac{(3a-1)^2}{12(b-a)(1-a-b)} \\ &\geq \frac{(3a-1)^2}{12a(a-1)} \\ &\geq \frac{2}{3},\end{aligned}$$

where the last inequality follows from $\frac{(3a-1)^2}{12a(a-1)}$ being decreasing on $(-\infty,-1]$.

Case 1(d): $0 < a+T_a < b+T_a \leq 1$

We seek to lower bound

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{(a+2b-3)^2}{12(1-a-b)}.$$

Taking the derivative with respect to b, we obtain

$$\begin{aligned}\frac{\partial}{\partial b}\left[\frac{(a+2b-3)^2}{12(1-a-b)}\right] &= \frac{48(1-a-b)(a+2b-3)+12(a+2b-3)^2}{(12(1-a-b))^2} \\ &= \frac{(a+2b-3)(1-3a-2b)}{12(1-a-b)^2}.\end{aligned}$$

Thus, since

$$a+2b-3 \leq 0$$

and

$$1-3a-2b \geq 0$$

whenever $a \in [-1,0]$ and $b \in [-1,0]$, we have that $\dfrac{\text{GFT}}{\text{FB-GFT}}$ is decreasing with respect to b on $a \in [-1,0]$ and $b \in [-1,0]$. Hence,

$$\begin{aligned}\frac{\text{GFT}}{\text{FB-GFT}} &= \frac{(a+2b-3)^2}{12(1-a-b)} \\ &\geq \frac{(a-3)^2}{12(1-a)} \\ &\geq \frac{2}{3},\end{aligned}$$

since $\dfrac{(a-3)^2}{12(1-a)}$ is increasing on $[-1,0]$.

Case 1(e): $0 < a + T_a < 1 < b + T_a$

Since

$$a + T_a > 0$$

$$a > -1,$$

we have that

$$b + T_a > 1$$

$$b > \frac{a+1}{2} > 0,$$

so this case is impossible because $b \leq 0$ in Case 1.

Case 2: $a \leq 0 < b \leq 1$

Case 2(a): $b + T_a \leq 0$

Since $T_a > 0$ and $b > 0$, this case is impossible.

Case 2(b): $a + T_a \leq 0 < b + T_a \leq 1$

Since

$$a + T_a \leq 0$$

$$a \leq -1$$

and

$$b + T_a \leq 1,$$

we have that

$$b \leq \frac{a+1}{2} \leq 0,$$

so this case is impossible because $b > 0$ in Case 2.

Case 2(c): $a + T_a \leq 0 < 1 < b + T_a$

Observe that $b^3 - 3b^2 + 3b$ is increasing on $[0, 1]$ and that $\dfrac{(3a-1)^2}{4(3a^2-3a+1)}$ is decreasing on $(-\infty, -1]$.

Thus, we have

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{(3a-1)^2}{4(3a^2-3a+b^3-3b^2+3b)} \geq \frac{(3a-1)^2}{4(3a^2-3a+1)} \geq \frac{4}{7}.$$

Case 2(d): $0 < a + T_a < b + T_a \leq 1$

We seek to lower bound

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{(a+2b-3)^2(b-a)}{4(3a^2-3a+b^3-3b^2+3b)}.$$

Taking the derivative with respect to b, we obtain

$$\begin{aligned}
&\frac{\partial}{\partial b}\left[\frac{(a+2b-3)^2(b-a)}{4(3a^2-3a+b^3-3b^2+3b)}\right] \\
&= \frac{4(3a^2-3a+b^3-3b^2+3b)(4(a+2b-3)(b-a)+(a+2b-3)^2)}{(4(3a^2-3a+b^3-3b^2+3b))^2} \\
&- \frac{4(a+2b-3)^2(b-a)(3b^2-6b+3)}{(4(3a^2-3a+b^3-3b^2+3b))^2} \\
&= -\frac{12(3a^2-3a+b^3-3b^2+3b)(a+2b-3)(a-2b+1)}{(4(3a^2-3a+b^3-3b^2+3b))^2} \\
&- \frac{12(a+2b-3)^2(b-a)(b-1)^2}{(4(3a^2-3a+b^3-3b^2+3b))^2} \\
&= -\frac{12(a+2b-3)}{(4(3a^2-3a+b^3-3b^2+3b))^2} \\
&\{(3a^3+ab^3-3ab^2+9ab-6a^2b-2b^4+7b^3-9b^2-3a+3b) \\
&+(-ab^3+2b^4-7b^3-a^2b^2+5ab^2+8b^2+2a^2b-7ab-3b-a^2+3a)\} \\
&= -\frac{12(a+2b-3)}{(4(3a^2-3a+b^3-3b^2+3b))^2}(3a^3+2ab^2+2ab-4a^2b-b^2-a^2b^2-a^2).
\end{aligned}$$

Now, since each of the terms in the rightmost factor is nonpositive and $a+2b-3 \leq 0$ whenever $a \in (-1,0]$ and $b \in \left(0, \frac{a+1}{2}\right]$, we have that $\frac{(a+2b-3)^2(b-a)}{4(3a^2-3a+b^3-3b^2+3b)}$ is decreasing with respect to b on $a \in (-1,0]$ and $b \in \left(0, \frac{a+1}{2}\right]$. Hence,

$$\begin{aligned}
\frac{\text{GFT}}{\text{FB-GFT}} &\geq \frac{(2a-2)^2((1-a)/2)}{4(3a^2-3a+(a^3-3a^2+3a+7)/8)} \\
&= \frac{4(1-a)^3}{a^3+21a^2-21a+7}.
\end{aligned}$$

Taking the derivative, we obtain

$$\frac{d}{da}\left[\frac{4(1-a)^3}{a^3+21a^2-21a+7}\right] = \frac{-12(1-a)^2(a^3+21a^2-21a+7)-4(1-a)^3(3a^2+42a-21)}{(a^3+21a^2-21a+7)^2}$$
$$= -\frac{96(1-a)^2a^2}{(a^3+21a^2-21a+7)^2},$$

so $\dfrac{4(1-a)^3}{a^3+21a^2-21a+7}$ is decreasing on $[-1,0]$. Therefore,

$$\frac{\text{GFT}}{\text{FB-GFT}} \geq \frac{4}{7}.$$

Case 2(e): $0 < a + T_a < 1 < b + T_a$

Since $b^3 - 3b^2 + 3b$ is increasing on $[0,1]$ and $\dfrac{(1-a)^3}{2(3a^2-3a+1)}$ is decreasing on $[-1,0]$, we have that

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{(1-a)^3}{2(3a^2-3a+b^3-3b^2+3b)} \geq \frac{(1-a)^3}{2(3a^2-3a+1)} \geq \frac{1}{2}.$$

Case 3: $a \leq 0 < 1 < b$

Cases 3(a), (b), (d): $b + T_a \leq 1$

Since $b > 1$ and $T_a > 0$, this case is impossible.

Case 3(c): $a + T_a \leq 0 < 1 < b + T_a$

Since $\dfrac{(3a-1)^2}{4(3a^2-3a+1)}$ is decreasing on $(-\infty, -1]$, we have that

$$\begin{aligned}\frac{\text{GFT}}{\text{FB-GFT}} &= \frac{(3a-1)^2}{4(3a^2-3a+1)} \\ &\geq \frac{4}{7}.\end{aligned}$$

Case 3(e): $0 < a + T_a < 1 < b + T_a$

Since $\dfrac{(1-a)^3}{2(3a^2-3a+1)}$ is decreasing on $[-1, 0]$, we have that

$$\begin{aligned}\frac{\text{GFT}}{\text{FB-GFT}} &= \frac{(1-a)^3}{2(3a^2-3a+1)} \\ &\geq \frac{1}{2}.\end{aligned}$$

Case 4: $0 < a < b \leq 1$

Case 4(a): $b + T_a \leq 0$

Since $b > 0$ and $T_a > 0$, this case is impossible.

Cases 4(b), (c): $a + T_a \leq 0$

Since $a > 0$ and $T_a > 0$, this case is impossible.

Case 4(d): $0 < a + T_a < b + T_a \leq 1$

We seek to minimize

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{(a+2b-3)^2}{4(a^2+ab+b^2-3a-3b+3)}.$$

Taking the derivative with respect to b,

$$\begin{aligned}\frac{\partial}{\partial b}\left[\frac{(a+2b-3)^2}{4(a^2+ab+b^2-3a-3b+3)}\right] &= \frac{16(a^2+ab+b^2-3a-3b+3)(a+2b-3) - 4(a+2b-3)^3}{(4(a^2+ab+b^2-3a-3b+3))^2} \\ &= \frac{12(a+2b-3)(a-1)^2}{(4(a^2+ab+b^2-3a-3b+3))^2}.\end{aligned}$$

Now, since $a + 2b - 3 \leq 0$ on $b \in \left(a, \frac{a+1}{2}\right]$, we have that $\frac{(a+2b-3)^2}{4(a^2+ab+b^2-3a-3b+3)}$ is decreasing with respect to b on $a \in (0,1)$ and $b \in \left(a, \frac{a+1}{2}\right]$. Thus,

$$\begin{aligned}\frac{\text{GFT}}{\text{FB-GFT}} &\geq \frac{(2a-2)^2}{4a^2+2a(a+1)+(a^2+2a+1)-12a-6(a+1)+12} \\ &= \frac{(2a-2)^2}{7(a-1)^2} \\ &= \frac{4}{7}.\end{aligned}$$

Case 4(e): $0 < a + T_a < 1 < b + T_a$

We seek to lower bound

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{(1-a)^3}{2(b-a)(a^2+ab+b^2-3a-3b+3)}.$$

Taking the derivative of the denominator with respect to b, we obtain

$$\begin{aligned}\frac{\partial}{\partial b}[2(b-a)(a^2+ab+b^2-3a-3b+3)] &= 2(b-a)(2b+a-3)+2(a^2+ab+b^2-3a-3b+3)\\ &= 6(b-1)^2,\end{aligned}$$

so $2(b-a)(a^2+ab+b^2-3a-3b+3)$ is increasing with respect to b on $[0,1]$. Thus,

$$\begin{aligned}\frac{\text{GFT}}{\text{FB-GFT}} &= \frac{(1-a)^3}{2(b-a)(a^2+ab+b^2-3a-3b+3)}\\ &\geq \frac{(1-a)^3}{2(1-a)(a^2-2a+1)}\\ &= \frac{1}{2}.\end{aligned}$$

Case 5: $0 < a < 1 < b$

Cases 5(a), (b), (d): $b + T_a \leq 1$

Since $b > 1$ and $T_a > 0$, this case is impossible.

Case 5(c): $a + T_a \leq 0 < 1 < b + T_a$

Since $a > 0$ and $T_a > 0$, this case is impossible.

Case 5(e): $0 < a + T_a < 1 < b + T_a$

We have that

$$\frac{\text{GFT}}{\text{FB-GFT}} = \frac{1}{2}.$$

Case 6: $a \geq 1$

We have that $\text{GFT} = 0 = \text{FB-GFT}$.

E.1.2 Derivations for Table E.2 and E.3

Here, we provide the derivations of each of the integrals in Tables E.2 and E.3. Each derivation is indexed by its corresponding number in the tables.

1.

$$\begin{aligned}\frac{1}{b-a}\int_0^1\int_a^b (v-c)\,dc\,dv &= \frac{1}{b-a}\int_0^1 \left[vc-\frac{c^2}{2}\right]_a^b dv \\ &= \frac{1}{b-a}\int_0^1 \left(v(b-a)-\frac{b^2-a^2}{2}\right) dv \\ &= \frac{1-a-b}{2}\end{aligned}$$

2.

$$\begin{aligned}&\frac{1}{b-a}\int_0^b\int_a^v (v-c)\,dc\,dv + \frac{1}{b-a}\int_b^1\int_a^b (v-c)\,dc\,dv \\ &= \frac{1}{b-a}\int_0^b \left[vc-\frac{c^2}{2}\right]_a^v dv + \frac{1}{b-a}\int_b^1 \left[vc-\frac{c^2}{2}\right]_a^b dv \\ &= \frac{3a^2-3a+b^3-3b^2+3b}{6(b-a)}\end{aligned}$$

3.

$$\begin{aligned}\frac{1}{b-a}\int_0^1\int_a^v (v-c)\,dc\,dv &= \frac{1}{b-a}\int_0^1 \left[vc-\frac{c^2}{2}\right]_a^v dv \\ &= \frac{3a^2-3a+1}{6(b-a)}\end{aligned}$$

4.

$$\begin{aligned}\frac{1}{b-a}\int_a^b\int_c^1 (v-c)\,dv\,dc &= \frac{1}{b-a}\int_a^b \left[\frac{v^2}{2}-cv\right]_c^1 dc \\ &= \frac{1}{b-a}\int_a^b \left[\frac{1-c^2}{2}-c(1-c)\right] dc \\ &= \frac{a^2+ab+b^2-3a-3b+3}{6}\end{aligned}$$

5.

$$\begin{aligned}\frac{1}{b-a}\int_a^1\int_a^v (v-c)\,dc\,dv &= \frac{1}{b-a}\int_a^1 \left[vc-\frac{c^2}{2}\right]_a^v dv \\ &= \frac{1}{b-a}\int_a^1 \left(v(v-a)-\frac{v^2-a^2}{2}\right) dv \\ &= \frac{(1-a)^3}{6(b-a)}\end{aligned}$$

6.

$$\begin{aligned}&\frac{1}{b-a}\int_0^{b+T_a}\int_a^{v-T_a}(v-c)\,dc\,dv + \frac{1}{b-a}\int_{b+T_a}^1\int_a^b (v-c)\,dc\,dv \\ &= \frac{1}{b-a}\int_0^{b+T_a}\left(v(v-T_a-a)-\frac{(v-T_a)^2-a^2}{2}\right)dv + \frac{1}{b-a}\int_{b+T_a}^1\left(v(b-a)-\frac{b^2-a^2}{2}\right)dv \\ &= \frac{1}{b-a}\int_0^{b+T_a}\left(\frac{v^2}{2}-av+\frac{3a^2+2a-1}{8}\right)dv + \int_{b+T_a}^1\left(v-\frac{a+b}{2}\right)dv \\ &= \frac{b+T_a}{b-a}\left(\frac{(b+T_a)^2}{6}-\frac{a(b+T_a)}{2}+\frac{3a^2+2a-1}{8}\right)+\left(\frac{1-(b+T_a)^2}{2}-\frac{(a+b)(1-b-T_a)}{2}\right) \\ &= \frac{(a+b-1)(a^2-4ab+10a+4b^2-8b+1)}{24(b-a)}\end{aligned}$$

7.

$$\begin{aligned}
\frac{1}{b-a}\int_0^1 \int_a^{v-T_a} (v-c)\,dc\,dv &= \frac{1}{b-a}\int_0^1 \left[vc - \frac{c^2}{2}\right]_a^{v-T_a} dv \\
&= \frac{1}{b-a}\int_0^1 \left(\frac{v^2}{2} - av - \frac{T_a^2}{2} + \frac{a^2}{2}\right) dv \\
&= \frac{1}{b-a}\int_0^1 \left(\frac{v^2}{2} - av + \frac{3a^2+2a-1}{8}\right) dv \\
&= \frac{(3a-1)^2}{24(b-a)}
\end{aligned}$$

8.

$$\begin{aligned}
\frac{1}{b-a}\int_a^b \int_{c+T_a}^1 (v-c)\,dv\,dc &= \frac{1}{b-a}\int_a^b \left(\frac{1-(c+T_a)^2}{2} - c(1-c-T_a)\right) dc \\
&= \frac{1}{b-a}\int_a^b \left(\frac{c^2}{2} - c + \frac{1-T_a^2}{2}\right) dc \\
&= \frac{1}{b-a}\left(\frac{b^3-a^3}{6} - \frac{b^2-a^2}{2} + \frac{(1-T_a^2)(b-a)}{2}\right) \\
&= \frac{a^2+ab+b^2}{6} - \frac{a+b}{2} + \frac{3+2a-a^2}{8} \\
&= \frac{(a+2b-3)^2}{24}
\end{aligned}$$

	Subcase				
Case	(a)	(b)	(c)	(d)	(e)
1	$[-3,-2]$	$[-1,0]$	$[-1-\epsilon,0]$	$[-1+\epsilon,0]$	
2			$[-1,1]$	$[0,1/2]$	$[0,1]$
3			$[-1,2]$		$[0,2]$
4				$[1/3,2/3]$	$[1/2,1]$
5					$[1/2,3/2]$

Table E.4: Examples of Values of a and b that Match Lower Bounds

9.

$$\begin{aligned}
\frac{1}{b-a}\int_{a+T_a}^{1}\int_{a}^{v-T_a}(v-c)\,dc\,dv &= \frac{1}{b-a}\int_{a+T_a}^{1}\left(\frac{v^2}{2}-av-\frac{T_a^2}{2}+\frac{a^2}{2}\right)dv \\
&= \frac{1}{b-a}\int_{a+T_a}^{1}\left(\frac{v^2}{2}-av+\frac{3a^2+2a-1}{8}\right)dv \\
&= \frac{(3a-1)^2}{24(b-a)}-\frac{(a+T_a)(a^2-4aT_a+4T_a^2+6a-3)}{24(b-a)} \\
&= \frac{(3a-1)^2}{24(b-a)}-\frac{(a+1)(2a^2+a-1)}{24(b-a)} \\
&= \frac{(1-a)^3}{12(b-a)}
\end{aligned}$$

E.1.3 Upper Bounds

Table E.4 summarizes the specific problem instances that yield the tightness results. In particular, $1/2$ is a tight lower bound on the approximation factor to the first-best GFT under the broker's optimal mechanism in the uniform setting. We omit the calculations as it is a simple plugging-in of the described values.

E.2 Proof of Theorem 6.4.6; Public-Buyer

We first introduce some notation:

Definition E.2.1. Given a direct mechanism MEC $= (x(c), p(c))$[1], we define the following functions for the expected utilities of the buyer, intermediary, and seller, respectively, assuming truthful reporting by the seller:

- $u_b = \int_a^b (x(c) \cdot v - p_b(c))\, G'(c)\, dc$
- $u_i = \int_a^b (p_b(c) - p_s(c))\, G'(c)\, dc$
- $u_s(c) = p_s(c) - x(c) \cdot c$

Note that with a public buyer, MEC is BNIC if $u_s(c) \geq p_s(\tilde{c}) - x(\tilde{c}) \cdot c$ for every c and $\tilde{c}$, and MEC is IR if $u_b \geq 0$ and $u_s(c) \geq 0$ for all c.

The following theorem is an analogue of Theorem 3 in [14]:

Theorem E.2.2. *For any BNIC mechanism* MEC *in the public-buyer setting,* $u_s(c)$ *is decreasing and*

$$u_i + u_b + u_s(b) = \int_a^b \left(v - \left(c + \frac{G(c)}{G'(c)}\right)\right) x(c)\, G'(c)\, dc.$$

Proof. Since MEC is BNIC, for every c and c',

$$u_s(c) = p_s(c) - x(c) \cdot c \geq p_s(c') - x(c') \cdot c$$

[1] In the public-buyer setting, the allocation function x and payment function p of MEC only depend on the seller's reported valuation c.

and

$$u_s(c') = p_s(c') - x(c') \cdot c' \geq p_s(c) - x(c) \cdot c'.$$

Thus,

$$(c' - c)x(c) \geq u_s(c) - u_s(c') \geq (c' - c)x(c').$$

In particular, if $c' > c$, then $x(c) \geq x(c')$, so $x(c)$ is decreasing and hence integrable. Therefore, $u_s'(c) = -x(c)$ almost everywhere, so

$$u_s(c) = u_s(b) + \int_v^b x(t)\,dt, \tag{E.1}$$

and $u_s(c)$ is decreasing.

Now, by (E.1), we have that

$$\begin{aligned}
\int_a^b (v - c)x(c)\,G'(c)\,dc - u_i &= \int_a^b (vx(c) - p_b(c))\,G'(c)\,dc + \int_a^b (p_s(c) - cx(c))\,G'(c)\,dc \\
&= u_b + \int_a^b u_s(c)\,G'(c)\,dc \\
&= u_b + \int_a^b \left(u_s(b) + \int_v^b x(t)\,dt\right) G'(c)\,dc \\
&= u_b + u_s(b) + \int_a^b \int_v^b x(t)\,dt\,G'(c)\,dc \\
&= u_b + u_s(b) + \int_a^b \int_a^t G'(c)\,dc\,x(t)\,dt \\
&= u_b + u_s(b) + \int_a^b G(t)\,x(t)\,dt.
\end{aligned}$$

Finally, from the first and last expressions in this chain of equations, we obtain

$$u_i + u_b + u_s(b) = \int_a^b \left(v - \left(c + \frac{G(c)}{G'(c)} \right) \right) x(c)\, G'(c)\, dc.$$

a

Next, the following is an analogue of Theorem 4 in [14], tailored to the public-buyer setting:

Theorem E.2.3. *Suppose G is a regular distribution in the public-buyer setting. Then, among all BNIC and IR mechanisms, the broker's expected profit is maximized by a mechanism in which the object is transferred to the buyer if and only if $\phi_G(c) \leq v$.*

Proof. By Theorem E.2.2, we have that

$$u_i = \int_a^b \left(v - \left(c + \frac{G(c)}{G'(c)} \right) \right) x(c)\, G'(c) - u_b - u_s(b).$$

Since we want an IR mechanism MEC that maximizes u_i, we want MEC to satisfy the following constraints:

1. $u_b = 0$
2. $u_s(b) = 0$
3. $x(c) = \begin{cases} 1 & \text{if } \phi_G(c) \leq v \\ 0 & \text{otherwise} \end{cases}$

It remains to define $p(c) = (p_s(c), p_b(c))$ such that first two constraints are satisfied. We take

$$p_s(c) = x(c) \cdot \sup\{t \leq b | \phi_G(t) \leq v\}$$

and

$$p_b(c) = x(c)v.$$

Then, since $x(c)v - p_b(c) = 0$ for all c and $p_s(b) = x(b)b$, we have that $u_b = 0$ and $u_s(b) = 0$, as desired.

Note that MEC is BNIC, since if trade occurs, the seller pays the highest valuation he could have reported and still have the trade occur. Then, since $u_s(c)$ is decreasing by Theorem E.2.2 and $u_s(b) = 0$, we have that MEC is IR. a

Theorem 6.4.6 directly follows from the above theorem.

Appendix F: Single-Sample Bilateral Trade with a Broker

F.1 Remaining Proofs for Preliminary Lemmas

F.1.1 Proof of Lemma 7.3.2

To prove Lemma 7.3.2, we will use the following result from [148]:

Lemma F.1.1. *[148] The gains-from-trade of a mechanism that offers price p to the buyer and price q to the seller, where $p \geq q$, is given by*

$$\text{GFT}(p,q) = \textsc{Profit}(p,q) + G(q)\int_p^\infty (1-F(x))\,dx + (1-F(p))\int_0^q G(x)\,dx.$$

Taking the expectation of this expression, we obtain the gains-from-trade of our mechanism as follows:

Proof of Lemma 7.3.2. By Lemma F.1.1, we have that

$$\text{GFT}(p,q) = \textsc{Profit}(p,q) + G(q)\int_p^\infty (1-F(x))\,dx + (1-F(p))\int_0^q G(x)\,dx$$

for $p \geq q$. Thus, together with Lemma 7.3.1, we have that

$$\mathbb{E}_{p,q\sim F,G}\left[\text{GFT}(p,q)\right] = \frac{1}{4}\int_0^\infty G(x)^2(1-F(x))^2\,dx + \lambda_1 + \lambda_2,$$

where we define

$$\lambda_1 = \int_0^\infty \int_q^\infty G(q) \int_p^\infty (1 - F(x))\, dx\, dF(p)\, dG(q)$$
$$\lambda_2 = \int_0^\infty \int_q^\infty (1 - F(p)) \int_0^q G(x)\, dx\, dF(p)\, dG(q).$$

Now, let $\lambda = \int_0^\infty (1 - F(x))\, dx$. Then, we have that

$$\begin{aligned}
\lambda_1 &= \int_0^\infty G(q) \left(\int_q^\infty \int_p^\infty (1 - F(x))\, dx\, dF(p) \right) dG(q) \\
&= \int_0^\infty G(q) \left(\int_q^\infty \left[\lambda - \int_0^p (1 - F(x))\, dx \right] dF(p) \right) dG(q) \\
&= \int_0^\infty G(q) \left(\lambda(1 - F(q)) - \int_q^\infty \int_0^p (1 - F(x))\, dx\, dF(p) \right) dG(q).
\end{aligned}$$

For the inner integral, notice that

$$\begin{aligned}
\int_q^\infty \int_0^p (1 - F(x))\, dx\, dF(p) &= F(p) \int_0^p (1 - F(x))\, dx \Big|_q^\infty - \int_q^\infty F(p)(1 - F(p))\, dp \\
&\qquad \text{(integrating by parts over } F'(p) \text{ and } \int_0^p (1 - F(x))\, dx) \\
&= \lambda - F(q) \int_0^q (1 - F(x))\, dx - \int_q^\infty F(p)(1 - F(p))\, dp.
\end{aligned}$$

Hence, we can rewrite λ_1 as

$$\begin{aligned}
\lambda_1 &= \int_0^\infty G(q)\left(F(q)\int_0^q (1-F(x))\,dx + \int_q^\infty F(p)(1-F(p))\,dp - \lambda F(q)\right) dG(q) \\
&= \int_0^\infty G(q)\left(\int_q^\infty (1-F(x))(F(x)-F(q))\,dx\right) dG(q) \\
&= \int_0^\infty G(q)\left(\int_q^\infty (1-F(x))F(x)\,dx\right) dG(q) - \int_0^\infty F(q)G(q)\left(\int_q^\infty (1-F(x))\,dx\right) dG(q) \\
&= \int_0^\infty G(q)\left(\int_q^\infty (1-F(x))F(x)\,dx\right) G'(q)\,dq - \int_0^\infty F(q)G(q)\left(\int_q^\infty (1-F(x))\,dx\right) G'(q)\,dq.
\end{aligned}$$

Integrating by parts again, we obtain

$$\begin{aligned}
&\lambda_1 \\
&= \left(\frac{G(q)^2}{2}\int_q^\infty (1-F(x))F(x)\,dx\bigg|_0^\infty + \frac{1}{2}\int_0^\infty G(q)^2F(q)(1-F(q))\,dq\right) \\
&\quad - \left(\frac{G(q)^2}{2}F(q)\int_q^\infty (1-F(x))\,dx\bigg|_0^\infty + \frac{1}{2}\int_0^\infty G(q)^2\left[F(q)(1-F(q)) - F'(q)\int_q^\infty (1-F(x))\,dx\right]\,dq\right) \\
&= \frac{1}{2}\int_0^\infty G(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q).
\end{aligned}$$

Finally, we have that

$$\begin{aligned}
\lambda_2 &= \int_0^\infty \int_q^\infty \left((1-F(p))\int_0^q G(x)\,dx\right) dF(p)\,dG(q) \\
&= \int_0^\infty \int_0^p \left((1-F(p))\int_0^q G(x)\,dx\right) dG(q)\,dF(p) \\
&= \int_0^\infty (1-F(p))\left(\int_0^p \int_0^q G(x)\,dx\,dG(q)\right) dF(p) \\
&= \int_0^\infty (1-F(p))\left(\int_0^p \int_0^q G(x)\,dx\,G'(q)\,dq\right) dF(p).
\end{aligned}$$

Integrating by parts over $G'(q)$ and $\int_0^q G(x)\,dx$, we obtain

$$
\begin{aligned}
&\lambda_2 \\
&= \int_0^\infty (1-F(p))\left(G(q)\int_0^q G(x)\,dx\Big|_0^p - \int_0^p G(q)^2\,dq\right) dF(p) \\
&= \int_0^\infty (1-F(p))\left(\int_0^p G(x)(G(p)-G(x))\,dx\right) dF(p) \\
&= \int_0^\infty (1-F(p))\left(\int_0^p G(x)(1-G(x))\,dx\right) dF(p) - \int_0^\infty (1-F(p))(1-G(p))\left(\int_0^p G(x)\,dx\right) dF(p) \\
&= \int_0^\infty (1-F(p))\left(\int_0^p G(x)(1-G(x))\,dx\right) F'(p)\,dp - \int_0^\infty (1-F(p))(1-G(p))\left(\int_0^p G(x)\,dx\right) F'(p)\,dp.
\end{aligned}
$$

Integrating by parts again, we obtain

$$
\begin{aligned}
\lambda_2 = &\left(-\frac{(1-F(p))^2}{2}\int_0^p G(x)(1-G(x))\,dx\Big|_0^\infty + \frac{1}{2}\int_0^\infty (1-F(p))^2(1-G(p))G(p)\,dp\right) \\
&- \left(-\frac{(1-F(p))^2}{2}(1-G(p))\int_0^p G(x)\,dx\Big|_0^\infty \right. \\
&\quad \left. + \frac{1}{2}\int_0^\infty (1-F(p))^2\left[(1-G(p))G(p) - G'(p)\int_0^p G(x)\,dx\right]\,dp\right) \\
&= \frac{1}{2}\int_0^\infty (1-F(p))^2\int_0^p G(x)\,dx\,dG(p).
\end{aligned}
$$

Finally, combining all our results from above, we conclude that

$$
\begin{aligned}
\mathbb{E}_{p,q\sim F,G}\left[\mathbf{GFT}(p,q)\right] = &\frac{1}{4}\int_0^\infty G(x)^2(1-F(x))^2\,dx + \frac{1}{2}\int_0^\infty G(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q) \\
&+ \frac{1}{2}\int_0^\infty (1-F(p))^2\int_0^p G(x)\,dx\,dG(p).
\end{aligned}
$$

a

F.1.2 Proof of Lemma 7.5.1

Proof of Lemma 7.5.1. By Lemma 7.3.2, we have that

$$\begin{aligned}\text{GFT} &= \mathbb{E}_{p,q\sim F,G}\left[\text{GFT}(p,q)\right] \\ &= \frac{1}{4}\int_0^\infty G(x)^2(1-F(x))^2\,dx + \frac{1}{2}\int_0^\infty G(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q) \\ &+ \frac{1}{2}\int_0^\infty (1-F(p))^2\int_0^p G(x)\,dx\,dG(p).\end{aligned}$$

Now, since F first-order stochastically dominates G, we have that

$$\begin{aligned}\int_0^\infty G(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q) &\geq \int_0^\infty F(q)^2\int_q^\infty (1-F(x))\,dx\,dF(q) \\ &= \frac{1}{3}\int_0^\infty F(x)^3(1-F(x))\,dx,\end{aligned}$$

where the last equality holds as in the proof of Lemma 7.4.1. Similarly, we have that

$$\begin{aligned}\int_0^\infty (1-F(p))^2\int_0^p G(x)\,dx\,dG(p) &\geq \int_0^\infty (1-G(p))^2\int_0^p G(x)\,dx\,dG(p) \\ &= \frac{1}{3}\int_0^\infty (1-G(x))^3G(x)\,dx.\end{aligned}$$

Thus,

$$\text{GFT} \geq \frac{1}{12}\int_0^\infty (3G(x)^2(1-F(x))^2 + 2F(x)^3(1-F(x)) + 2(1-G(x))^3G(x))\,dx.$$

a

F.2 Preliminaries for Upper-Bound Proofs

When proving upper bounds on approximation ratios, we will often use distributions H such that $H(x) \approx \alpha$ for some constant $0 < \alpha < 1$. To construct such distributions, we first recall the following definition:

Definition F.2.1. Let $0 < \alpha < 1$ and $0 < \delta < \dfrac{\alpha}{2\alpha+1}$. Define $H_{\alpha,\delta,\text{QUAD}} : [0,\delta] \to \mathbb{R}$ as

$$H_{\alpha,\delta,\text{QUAD}}(x) = \frac{\alpha-\delta}{\delta}x + \frac{(2\alpha+1)\delta-\alpha}{\delta^2(1-2\delta)}x(x-\delta)$$

and $H_{\alpha,\delta,\text{LIN}} : [\delta, 1-\delta] \to \mathbb{R}$ as

$$H_{\alpha,\delta,\text{LIN}}(x) = \alpha + \frac{2\delta}{1-2\delta}\left(x - \frac{1}{2}\right).$$

We now recall the definition for our desired class of distributions:

Definition F.2.2. Let $0 < \alpha < 1$ and $0 < \delta < \min\left(\dfrac{\alpha}{2\alpha+1}, \dfrac{1-\alpha}{3-2\alpha}\right)$. We define $H_{\alpha,\delta} : [0,1] \to \mathbb{R}$ as

$$H_{\alpha,\delta}(x) = \begin{cases} H_{\alpha,\delta,\text{QUAD}}(x) & \text{if } 0 \le x \le \delta \\ H_{\alpha,\delta,\text{LIN}}(x) & \text{if } \delta \le x \le 1-\delta \\ 1 - H_{1-\alpha,\delta,\text{QUAD}}(1-x) & \text{if } 1-\delta \le x \le 1. \end{cases}$$

We first show that $H_{\alpha,\delta}$ is well-defined and differentiable:

Lemma F.2.3. *Let* $0 < \alpha < 1$ *and* $0 < \delta < \min\left(\dfrac{\alpha}{2\alpha+1}, \dfrac{1-\alpha}{3-2\alpha}\right)$. *Then,* $H_{\alpha,\delta}$ *is well-defined and*

differentiable.

Proof. Note that $H_{1-\alpha,\delta,\text{QUAD}}(1-x)$ is well-defined on $[1-\delta, 1]$, since $\delta < \dfrac{1-\alpha}{3-2\alpha} = \dfrac{1-\alpha}{2(1-\alpha)+1}$. In addition, note the following:

1. $H_{\alpha,\delta,\text{QUAD}}(x)$ is the unique quadratic satisfying $H_{\alpha,\delta,\text{QUAD}}(0) = 0, H_{\alpha,\delta,\text{QUAD}}(\delta) = \alpha - \delta$, and $H'_{\alpha,\delta,\text{QUAD}}(\delta) = 2\delta/(1-2\delta)$.

2. $H_{\alpha,\delta,\text{LIN}}(x)$ is the unique linear function $H_{\alpha,\delta,\text{LIN}}(\delta) = \alpha - \delta, H_{\alpha,\delta,\text{LIN}}(1-\delta) = \alpha + \delta$, and $H'_{\alpha,\delta,\text{LIN}}(\delta) = 2\delta/(1-2\delta) = H'_{\alpha,\delta,\text{LIN}}(1-\delta)$.

3. $Q(x) = 1 - H_{1-\alpha,\delta,\text{QUAD}}(1-x)$ is the unique quadratic satisfying $Q(1) = 1, Q(1-\delta) = \alpha + \delta$, and $Q'(1-\delta) = 2\delta/(1-2\delta)$.

Thus, since the different components and their derivatives match at the boundary points δ and $1-\delta$, we have that $H_{\alpha,\delta}$ is well-defined and differentiable. a

Now, we want to show that $H_{\alpha,\delta}$ is a valid CDF. However, it is unclear whether $H_{\alpha,\delta,\text{QUAD}}$ is strictly increasing. The following lemma verifies this:

Lemma F.2.4. *Let* $0 < \alpha < 1$ *and* $0 < \delta < \dfrac{\alpha}{2\alpha+1}$. *Then,* $H_{\alpha,\delta,QUAD}$ *is strictly increasing.*

Proof. Taking the derivative of $H_{\alpha,\delta,\text{QUAD}}$, we obtain

$$H'_{\alpha,\delta,\text{QUAD}}(x) = \frac{\alpha-\delta}{\delta} + \frac{(2\alpha+1)\delta - \alpha}{\delta^2(1-2\delta)}(2x-\delta).$$

Now, $H'_{\alpha,\delta,\text{QUAD}}(x) > 0$ is equivalent to

$$2x - \delta < \frac{\delta(1-2\delta)(\delta-\alpha)}{(2\alpha+1)\delta - \alpha}.$$

(Note that $(2\alpha+1)\delta - \alpha < 0$ and $1 - 2\delta > 0$, since $\delta < \alpha/(2\alpha+1) < 1/2$.). Rewriting, we obtain

$$x < \delta\left(1 - \frac{\delta^2}{(2\alpha+1)\delta - \alpha}\right),$$

so it suffices if $x < \delta$, which is true on the domain of $H_{\alpha,\delta,\text{QUAD}}$. Thus, $H_{\alpha,\delta,\text{QUAD}}$ is strictly increasing.

a

Finally, we combine our results above to show that $H_{\alpha,\delta}$ is a valid CDF:

Lemma F.2.5. *Let* $0 < \alpha < 1$ *and* $0 < \delta < \min\left(\frac{\alpha}{2\alpha+1}, \frac{1-\alpha}{3-2\alpha}\right)$. *Then,* $H_{\alpha,\delta}$ *is a CDF that is differentiable everywhere on its domain.*

Proof. By Lemma F.2.3, $H_{\alpha,\delta}$ is well-defined and differentiable. In addition, $H_{\alpha,\delta}(0) = 0$ and $H_{\alpha,\delta}(1) = 1$. Thus, to show that $H_{\alpha,\delta}$ is a valid CDF, and it suffices to show that $H_{\alpha,\delta}$ is strictly increasing. Clearly, $H_{\alpha,\delta,\text{LIN}}$ is strictly increasing, and by Lemma F.2.4, $H_{\alpha,\delta,\text{QUAD}}$ is strictly increasing on $[0,\delta]$ and $1 - H_{1-\alpha,\delta,\text{QUAD}}$ is strictly increasing on $[1-\delta, 1]$, which completes the proof. a

Now, since we will apply the above class of distributions to the stochastic dominance setting, the following lemma shows that the family of distributions $H_{\alpha,\delta}$ is stochastically ordered with respect to α if δ is sufficiently small.

Lemma F.2.6. *Let* $0 < \alpha_F \leq \alpha_G < 1$, *and let*

$$0 < \delta < \min\left(\frac{\alpha_F}{2\alpha_F+1}, \frac{1-\alpha_F}{3-2\alpha_F}, \frac{\alpha_G}{2\alpha_G+1}, \frac{1-\alpha_G}{3-2\alpha_G}\right).$$

Then, $F = H_{\alpha_F,\delta}$ *stochastically dominates* $G = H_{\alpha_G,\delta}$.

Proof. We have that

$$\begin{aligned} H_{\alpha_G,\delta}(x) &= \frac{\alpha_G - \delta}{\delta}x + \frac{(2\delta-1)\alpha_G + \delta}{\delta^2(1-2\delta)}x(x-\delta) \\ &\geq \frac{\alpha_F - \delta}{\delta}x + \frac{(2\delta-1)\alpha_F + \delta}{\delta^2(1-2\delta)}x(x-\delta) \\ &= H_{\alpha_F,\delta}(x) \end{aligned}$$

for all $0 \leq x \leq \delta$, where the inequality holds since $2\delta - 1 < 0$ and $x - \delta \leq 0$. Thus, $G(x) \geq F(x)$ on $[0, \delta]$, and letting $1 - \alpha_G$ and $1 - \alpha_F$ take the roles of α_F and α_G, respectively, in the above inequality, we obtain that

$$G(x) = 1 - H_{1-\alpha_G}(1-x) \geq 1 - H_{1-\alpha_F}(1-x) = F(x)$$

on $[1 - \delta, 1]$. Finally, we have that

$$G(x) = F(x) + (\alpha_G - \alpha_F) \geq F(x)$$

on $[\delta, 1 - \delta]$. a

Next, we seek to upper-bound the gains-from-trade of our mechanism for $F = H_{\alpha_F,\delta}$ and $G = H_{\alpha_G,\delta}$ as $\delta \to 0$. The following lemma will be helpful in this regard:

Lemma F.2.7. *Let* $0 < \alpha_F \leq \alpha_G < 1$. *Then,*

$$\int_0^\delta (H_{\alpha_G,\delta,\textit{QUAD}}(x)^2 - H_{\alpha_F,\delta,\textit{QUAD}}(x)^2) H'_{\alpha_F,\delta,\textit{QUAD}}(x)\, dx = \frac{(\alpha_G - \alpha_F)(\alpha_G + \alpha_F)\alpha_F}{3} + O(\delta),$$

where the big-O notation hides dominated terms with respect to δ.

Proof. For $0 < \alpha < 1$, we have that

$$\begin{aligned} H_{\alpha,\delta,\text{QUAD}}(x) &= \frac{\alpha-\delta}{\delta}x + \frac{(2\alpha+1)\delta-\alpha}{\delta^2(1-\delta)}x(x-\delta) \\ &= O\left(\frac{\alpha}{\delta}\right)x + \left(O\left(\frac{\alpha}{\delta}\right)x - O\left(\frac{\alpha}{\delta^2}\right)x^2\right) \\ &= O\left(\frac{2\alpha}{\delta}\right)x - O\left(\frac{\alpha}{\delta^2}\right)x^2 \end{aligned}$$

and that

$$\begin{aligned} H'_{\alpha,\delta,\text{QUAD}}(x) &= \frac{\alpha-\delta}{\delta} + \frac{(2\alpha+1)\delta-\alpha}{\delta^2(1-\delta)}(2x-\delta) \\ &= O\left(\frac{\alpha}{\delta}\right) + \left(O\left(\frac{\alpha}{\delta}\right) - O\left(\frac{2\alpha}{\delta^2}\right)x\right) \\ &= O\left(\frac{2\alpha}{\delta}\right) - O\left(\frac{2\alpha}{\delta^2}\right)x. \end{aligned}$$

Thus,

$$\begin{aligned} &\int_0^\delta (H_{\alpha_G,\delta,\text{QUAD}}(x)^2 - H_{\alpha_F,\delta,\text{QUAD}}(x)^2)H'_{\alpha_F,\delta,\text{QUAD}}(x)\,dx \\ &\quad= \int_0^\delta (H_{\alpha_G,\delta,\text{QUAD}}(x) - H_{\alpha_F,\delta,\text{QUAD}}(x))(H_{\alpha_G,\delta,\text{QUAD}}(x) + H_{\alpha_F,\delta,\text{QUAD}}(x))H'_{\alpha_F,\delta,\text{QUAD}}(x)\,dx \\ &\quad= \int_0^\delta \left(\left(O\left(\frac{2(\alpha_G-\alpha_F)}{\delta}\right)x - O\left(\frac{\alpha_G-\alpha_F}{\delta^2}\right)x^2\right)\left(O\left(\frac{2(\alpha_G+\alpha_F)}{\delta}\right)x - O\left(\frac{\alpha_G+\alpha_F}{\delta^2}\right)x^2\right)\right. \\ &\qquad\left.\left(O\left(\frac{2\alpha_F}{\delta}\right) - O\left(\frac{2\alpha_F}{\delta^2}\right)x\right)\,dx\right). \end{aligned}$$

Finally, expanding the integrand and collecting like terms, we obtain

$$
\begin{aligned}
&\int_0^\delta (H_{\alpha_G,\delta,\text{QUAD}}(x)^2 - H_{\alpha_F,\delta,\text{QUAD}}(x)^2) H'_{\alpha_F,\delta,\text{QUAD}}(x)\,dx \\
&= \int_0^\delta \left(O\left(\frac{8(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^3}\right)x^2 - O\left(\frac{16(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^4}\right)x^3 \right. \\
&\quad \left. + O\left(\frac{10(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^5}\right)x^4 - O\left(\frac{2(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^6}\right)x^5 \right) dx \\
&= \left[O\left(\frac{8(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^3}\right)\frac{x^3}{3} - O\left(\frac{16(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^4}\right)\frac{x^4}{4} \right. \\
&\quad \left. + O\left(\frac{10(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^5}\right)\frac{x^5}{5} - O\left(\frac{2(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{\delta^6}\right)\frac{x^6}{6} \right]_0^\delta \\
&= \frac{8(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{3} - \frac{16(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{4} + \frac{10(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{5} \\
&\quad - \frac{2(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{6} + O(\delta) \\
&= \frac{(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{3} + O(\delta).
\end{aligned}
$$

a

We now upper-bound the gains-from-trade of our mechanism for $F = H_{\alpha_F,\delta}$ and $G = H_{\alpha_G,\delta}$ as $\delta \to 0$ in the following lemma:

Lemma F.2.8. *Let* $0 < \alpha_F \le \alpha_G < 1$, *Then, for any* ϵ, *there exists*

$$
0 < \delta < \min\left(\frac{\alpha_F}{2\alpha_F+1}, \frac{1-\alpha_F}{3-2\alpha_F}, \frac{\alpha_G}{2\alpha_G+1}, \frac{1-\alpha_G}{3-2\alpha_G}\right)
$$

such that if $F = H_{\alpha_F,\delta}$ *and* $G = H_{\alpha_G,\delta}$, *the gains-from-trade of a mechanism that draws samples*

$p \sim F, q \sim G$, and offers price p to the buyer and price q to the seller achieves

$$\text{GFT} < \frac{1}{12}\alpha_G(1-\alpha_F)(\alpha_G - 2\alpha_F + \alpha_F\alpha_G + 2) + \epsilon.$$

Proof. By Lemma 7.3.2, we have that

$$\begin{aligned}\text{GFT} = \frac{1}{4}\int_0^1 G(x)^2(1-F(x))^2\,dx + \frac{1}{2}\int_0^1 G(q)^2\int_q^1 (1-F(x))\,dx\,dF(q) \\ + \frac{1}{2}\int_0^1 (1-F(p))^2\int_0^p G(x)\,dx\,dG(p). \end{aligned} \tag{F.1}$$

We now analyze each of the integrals in (F.1).

First Integral in (F.1). We have that

$$\begin{aligned}\int_0^1 G(x)^2(1-F(x))^2\,dx &= \left(\int_0^\delta G(x)^2(1-F(x))^2\,dx + \int_\delta^{1-\delta} G(x)^2(1-F(x))^2\,dx \right. \\ &\qquad \left. + \int_0^{1-\delta} G(x)^2(1-F(x))^2\,dx\right) \\ &\le \left(2\delta + \int_\delta^{1-\delta} G(x)^2(1-F(x))^2\,dx\right) \\ &\le \left(2\delta + \int_\delta^{1-\delta} (\alpha_G+\delta)^2(1-(\alpha_F-\delta))^2\,dx\right) \\ &= 2\delta + (1-2\delta)(\alpha_G+\delta)^2(1-(\alpha_F-\delta))^2,\end{aligned}$$

where the last inequality holds since $F(x) \ge \alpha_F - \delta$ and $G(x) \le \alpha_G + \delta$ on $[\delta, 1-\delta]$.

Second Integral in (F.1). We have that

$$\begin{aligned}
&\int_0^1 G(q)^2 \int_q^1 (1-F(x))\,dx\,dF(q) \\
&\quad= \int_0^1 F(q)^2 \int_q^1 (1-F(x))\,dx\,dF(q) + \int_0^1 (G(q)^2 - F(q)^2) \int_q^1 (1-F(x))\,dx\,dF(q) \\
&\quad= \underbrace{\int_0^1 F(q)^2 \int_q^1 (1-F(x))\,dx\,dF(q)}_{(a)} + \underbrace{\int_0^\delta (G(q)^2 - F(q)^2) \int_q^1 (1-F(x))\,dx\,dF(q)}_{(b)} \\
&\qquad+ \underbrace{\int_\delta^{1-\delta} (G(q)^2 - F(q)^2) \int_q^1 (1-F(x))\,dx\,dF(q)}_{(c)} + \underbrace{\int_{1-\delta}^1 (G(q)^2 - F(q)^2) \int_q^1 (1-F(x))\,dx\,dF(q)}_{(d)}.
\end{aligned}$$

Now, integrating by parts as in the proof of 7.4.1, we obtain

$$\begin{aligned}
(a) &= \frac{1}{3}\int_0^1 F(x)^3(1-F(x))\,dx \\
&= \frac{1}{3}\left(\int_0^\delta F(x)^3(1-F(x))\,dx + \int_\delta^{1-\delta} F(x)^3(1-F(x))\,dx + \int_{1-\delta}^1 F(x)^3(1-F(x))\,dx\right) \\
&\le \frac{1}{3}\left(\delta + \int_\delta^{1-\delta} (\alpha_F+\delta)^3(1-(\alpha_F-\delta))\,dx + \delta\right) \\
&= \frac{1}{3}(2\delta + (1-2\delta)(\alpha_F+\delta)^3(1-(\alpha_F-\delta))).
\end{aligned}$$

In addition,

$$\begin{aligned}
(b) &= \int_0^\delta (G(q)^2 - F(q)^2) \left(\int_q^\delta (1 - F(x))\, dx + \int_\delta^1 (1 - F(x))\, dx \right) dF(q) \\
&= \int_0^\delta (G(q)^2 - F(q)^2) \left(\delta + \int_\delta^1 (1 - (\alpha_F - \delta))\, dx \right) dF(q) \\
&\le \int_0^\delta (G(q)^2 - F(q)^2)(\delta + (1-\delta)(1 - (\alpha_F - \delta)))\, dF(q) \\
&= (\delta + (1-\delta)(1 - (\alpha_F - \delta))) \int_0^\delta (G(q)^2 - F(q)^2)\, dF(q) \\
&= (\delta + (1-\delta)(1 - (\alpha_F - \delta))) \int_0^\delta (H_{\alpha_G,\delta,\mathrm{QUAD}}(q)^2 - H_{\alpha_F,\delta,\mathrm{QUAD}}(q)^2) H'_{\alpha_F,\delta,\mathrm{QUAD}}(q)\, dq \\
&= (\delta + (1-\delta)(1 - (\alpha_F - \delta))) \left(\frac{(\alpha_G - \alpha_F)(\alpha_G + \alpha_F)\alpha_F}{3} + O(\delta) \right),
\end{aligned}$$

where the last equality holds by Lemma F.2.7. Finally,

$$(c) \le \int_\delta^{1-\delta} (G(q)^2 - F(q)^2)\, dF(q) \le \int_\delta^{1-\delta} dF(q) = 2\delta$$

and

$$\begin{aligned}
(d) &\le \int_{1-\delta}^1 (G(q)^2 - F(q)^2) \int_{1-\delta}^1 (1 - F(x))\, dx\, dF(q) \\
&\le \delta \int_{1-\delta}^1 (G(q)^2 - F(q)^2)\, dF(q) \\
&\le \delta \int_{1-\delta}^1 dF(q) \\
&= \delta(1 - (\alpha_F + \delta)).
\end{aligned}$$

Combining our results from above, we obtain

$$\begin{aligned}\int_0^1 G(q)^2 \int_q^1 (1-F(x))\,dx\,dF(q) \leq \frac{1}{3}(2\delta + (1-2\delta)(\alpha_F+\delta)^3(1-(\alpha_F-\delta))) \\ + (\delta + (1-\delta)(1-(\alpha_F-\delta)))\left(\frac{(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{3} + O(\delta)\right) \\ + 2\delta + \delta(1-(\alpha_F+\delta)).\end{aligned}$$

Third Integral in (F.1). We have that

$$\begin{aligned}&\int_0^1 (1-F(p))^2 \int_0^p G(x)\,dx\,dG(p) \\ &= \int_0^1 (1-G(p))^2 \int_0^p G(x)\,dx\,dG(p) + \int_0^1 ((1-F(p))^2 - (1-G(p))^2) \int_0^p G(x)\,dx\,dG(p) \\ &= \underbrace{\int_0^1 (1-G(p))^2 \int_0^p G(x)\,dx\,dG(p)}_{(a)} \\ &\quad + \underbrace{\int_0^\delta ((1-F(p))^2 - (1-G(p))^2) \int_0^p G(x)\,dx\,dG(p)}_{(b)} \\ &\quad + \underbrace{\int_\delta^{1-\delta} ((1-F(p))^2 - (1-G(p))^2) \int_0^p G(x)\,dx\,dG(p)}_{(c)} \\ &\quad + \underbrace{\int_{1-\delta}^1 ((1-F(p))^2 - (1-G(p))^2) \int_0^p G(x)\,dx\,dG(p)}_{(d)}.\end{aligned}$$

Now, integrating by parts as in the proof of Lemma 7.4.1, we obtain

$$\begin{aligned}
(a) &= \frac{1}{3}\int_0^1 G(x)(1-G(x))^3\,dx \\
&= \frac{1}{3}\left(\int_0^\delta G(x)(1-G(x))^3\,dx + \int_\delta^{1-\delta} G(x)(1-G(x))^3\,dx + \int_{1-\delta}^1 G(x)(1-G(x))^3\,dx\right) \\
&\le \frac{1}{3}\left(\delta + \int_\delta^{1-\delta} (\alpha_G+\delta)(1-(\alpha_G-\delta))^3\,dx + \delta\right) \\
&= \frac{1}{3}(2\delta + (1-2\delta)(\alpha_G+\delta)(1-(\alpha_G-\delta))^3).
\end{aligned}$$

In addition, we have that

$$\begin{aligned}
(d) &= \int_{1-\delta}^1 ((1-F(p))^2 - (1-G(p))^2)\left(\int_0^{1-\delta} G(x)\,dx + \int_{1-\delta}^p G(x)\,dx\right)\,dG(p) \\
&\le \int_{1-\delta}^1 ((1-F(p))^2 - (1-G(p))^2)\left(\int_0^{1-\delta} (\alpha_G+\delta)\,dx + \delta\right)\,dG(p) \\
&= \int_{1-\delta}^1 ((1-F(p))^2 - (1-G(p))^2)((1-\delta)(\alpha_G+\delta)+\delta)\,dG(p) \\
&= ((1-\delta)(\alpha_G+\delta)+\delta)\int_{1-\delta}^1 ((1-F(p))^2 - (1-G(p))^2)\,dG(p) \\
&= ((1-\delta)(\alpha_G+\delta)+\delta)\int_{1-\delta}^1 (H_{1-\alpha_F,\delta,\text{QUAD}}(1-p)^2 - H_{1-\alpha_G,\delta,\text{QUAD}}(1-p)^2)H'_{1-\alpha_G,\delta,\text{QUAD}}(1-p)\,dp \\
&= ((1-\delta)(\alpha_G+\delta)+\delta)\int_0^{\delta} (H_{1-\alpha_F,\delta,\text{QUAD}}(p)^2 - H_{1-\alpha_G,\delta,\text{QUAD}}(p)^2)H'_{1-\alpha_G,\delta,\text{QUAD}}(p)\,dp \\
&= ((1-\delta)(\alpha_G+\delta)+\delta)\left(\frac{(\alpha_G-\alpha_F)(2-\alpha_F-\alpha_G)(1-\alpha_G)}{3} + O(\delta)\right),
\end{aligned}$$

where we perform a change of variables in the second-to-last equality and where the last equality holds

by Lemma F.2.7. Finally, we have that

$$\begin{aligned}(b) &\leq \int_0^\delta ((1-F(p))^2 - (1-G(p))^2) \int_0^\delta G(x)\,dx\,dG(p) \\ &\leq \delta \int_0^\delta ((1-F(p))^2 - (1-G(p))^2)\,dG(p) \\ &\leq \delta \int_0^\delta dG(p) \\ &= \delta(\alpha_G - \delta)\end{aligned}$$

and

$$(c) \leq \int_\delta^{1-\delta} ((1-F(p))^2 - (1-G(p))^2)\,dG(p) \leq \int_\delta^{1-\delta} dG(p) = 2\delta.$$

Combining our results from above, we obtain

$$\begin{aligned}\int_0^1 (1-F(p))^2 \int_0^p G(x)\,dx\,dG(p) \leq \frac{1}{3}&(2\delta + (1-2\delta)(\alpha_G+\delta)(1-(\alpha_G-\delta))^3) \\ &+ ((1-\delta)(\alpha_G+\delta)+\delta)\left(\frac{(\alpha_G-\alpha_F)(2-\alpha_F-\alpha_G)(1-\alpha_G)}{3} + O(\delta)\right) \\ &+ \delta(\alpha_G-\delta) + 2\delta.\end{aligned}$$

Finally, returning to (F.1), we obtain that for sufficiently small δ,

$$\begin{aligned}\text{GFT} &< \frac{\alpha_G^2(1-\alpha_F)^2}{4} + \frac{\alpha_F^3(1-\alpha_F)}{6} + \frac{(1-\alpha_F)(\alpha_G-\alpha_F)(\alpha_G+\alpha_F)\alpha_F}{6} \\ &\quad + \frac{\alpha_G(1-\alpha_G)^3}{6} + \frac{\alpha_G(\alpha_G-\alpha_F)(2-\alpha_F-\alpha_G)(1-\alpha_G)}{6} + \epsilon \\ &= \frac{1}{12}\alpha_G(1-\alpha_F)(\alpha_G - 2\alpha_F + \alpha_F\alpha_G + 2) + \epsilon.\end{aligned}$$

a

F.3 Remaining Proofs for Symmetric Agents

F.3.1 Proof of Theorem 7.4.3

Proof of Theorem 7.4.3. We have that

$$\begin{aligned}
\text{SW} &= \mathbb{E}_{c\sim F}\left[c\right] + \text{GFT} \\
&= \int_0^\infty (1-F(x))\,dx + \left(\frac{1}{3}\int_0^\infty F(x)(1-F(x))\,dx - \frac{1}{6}\int_0^\infty F(x)^2(1-F(x))^2\,dx\right) \\
&= \frac{1}{6}\left(\int_0^\infty (1-F(x))(1+F(x))(F(x)^2-2F(x)+4)\,dx + 2\int_0^\infty (1-F(x))\,dx\right),
\end{aligned}$$

where the second equality holds by Lemma 7.4.1. Thus, since $F(x)^2-2F(x)+4 = (F(x)-1)^2+3 \geq 3$ and $1+F(x) \leq 2$ for all x, we have that

$$\begin{aligned}
\text{SW} &= \frac{1}{6}\left(\int_0^\infty (1-F(x))(1+F(x))(F(x)^2-2F(x)+4)\,dx + 2\int_0^\infty (1-F(x))\,dx\right) \\
&\geq \frac{1}{6}\left(3\int_0^\infty (1-F(x))(1+F(x))\,dx + \int_0^\infty (1-F(x))(1+F(x))\,dx\right) \\
&= \frac{2}{3}\int_0^\infty (1-F(x))(1+F(x))\,dx \\
&= \frac{2}{3}\text{FB-SW}.
\end{aligned}$$

a

F.4 Remaining Proofs for Asymmetric Agents

F.4.1 Proof of Theorem 7.5.5

Proof of 7.5.5. Let $0 < \alpha_F \leq \alpha_G < 1$ be constants to be determined later, and let

$$0 < \delta < \min\left(\frac{\alpha_F}{2\alpha_F + 1}, \frac{1-\alpha_F}{3-2\alpha_F}, \frac{\alpha_G}{2\alpha_G+1}, \frac{1-\alpha_G}{3-2\alpha_G}\right).$$

Now, let $F = H_{\alpha_F,\delta}$ and $G = H_{\alpha_G,\delta}$. Note that F stochastically dominates G by Lemma F.2.6. Then, we have that

$$\begin{aligned}
\text{FB-GFT} &= \int_0^1 G(x)(1-F(x))\,dx \\
&\geq \int_\delta^{1-\delta} G(x)(1-F(x))\,dx \\
&\geq \int_\delta^{1-\delta} (\alpha_G - \delta)(1-(\alpha_F+\delta))\,dx \\
&= (1-2\delta)(\alpha_G-\delta)(1-(\alpha_F+\delta)).
\end{aligned}$$

Thus, by Lemma F.2.8, for sufficiently small δ, we have that

$$\frac{\text{GFT}}{\text{FB-GFT}} < \frac{1}{12}(\alpha_G - 2\alpha_F + \alpha_F\alpha_G + 2) + \epsilon.$$

Finally, since the above ratio is minimized when $\alpha_F = \alpha_G$ and $\alpha_G = 1/2$, we obtain that for sufficiently small δ,

$$\frac{\text{GFT}}{\text{FB-GFT}} < \frac{7}{48} + \epsilon.$$

a

F.4.2 Proof of Theorem 7.5.6

Proof of Theorem 7.5.6. Our proof follows a structure similar to that of the proof of Theorem 7.5.5. Let $0 < \alpha_F \leq \alpha_G < 1$ be constants to be determined later, and let

$$0 < \delta < \min\left(\frac{\alpha_F}{2\alpha_F + 1}, \frac{1-\alpha_F}{3-2\alpha_F}, \frac{\alpha_G}{2\alpha_G + 1}, \frac{1-\alpha_G}{3-2\alpha_G}\right).$$

Now, let $F = H_{\alpha_F,\delta}$ and $G = H_{\alpha_G,\delta}$. Note that F stochastically dominates G by Lemma F.2.6. Now, we have that

$$\begin{aligned}
\mathbb{E}_{c\sim G}[c] &= \int_0^1 (1 - G(x))\, dx \\
&\leq 2\delta + \int_\delta^{1-\delta} (1 - G(x))\, dx \\
&\leq 2\delta + \int_\delta^{1-\delta} (1 - (\alpha_G - \delta))\, dx \\
&= 2\delta + (1-2\delta)(1 - (\alpha_G - \delta))
\end{aligned}$$

and

$$\begin{aligned}
\text{FB-SW} &= \mathbb{E}_{c\sim G}\left[c\right] + \text{FB-GFT} \\
&= \int_0^1 (1 - G(x))\,dx + \int_0^1 G(x)(1 - F(x))\,dx \\
&\geq \int_\delta^{1-\delta} (1 - G(x))\,dx + \int_\delta^{1-\delta} G(x)(1 - F(x))\,dx \\
&\geq \int_\delta^{1-\delta} (1 - (\alpha_G + \delta))\,dx + \int_\delta^{1-\delta} (\alpha_G - \delta)(1 - (\alpha_F + \delta))\,dx \\
&= (1 - 2\delta)(1 - (\alpha_G - \delta)) + (1 - 2\delta)(\alpha_G - \delta)(1 - (\alpha_F + \delta)).
\end{aligned}$$

Thus, since $\text{GFT} = \mathbb{E}_{c\sim G}\left[c\right]$, by Lemma F.2.8, for sufficiently small δ, we have that

$$\begin{aligned}
\frac{\text{SW}}{\text{FB-SW}} &< \frac{12(1-\alpha_G) + \alpha_G(1-\alpha_F)(\alpha_G - 2\alpha_F + \alpha_F\alpha_G + 2)}{12(1-\alpha_F\alpha_G)} + \epsilon \\
&= \frac{12 - \alpha_F^2\alpha_G^2 + 2\alpha_F^2\alpha_G - 4\alpha_F\alpha_G + \alpha_G^2 - 10\alpha_G}{12(1-\alpha_F\alpha_G)} + \epsilon.
\end{aligned}$$

It remains to minimize the above ratio subject to $0 < \alpha_F \leq \alpha_G < 1$. Taking the derivative with respect to α_G, we obtain

$$\begin{aligned}
&\frac{\partial}{\partial\alpha_G}\left[\frac{12 - \alpha_F^2\alpha_G^2 + 2\alpha_F^2\alpha_G - 4\alpha_F\alpha_G + \alpha_G^2 - 10\alpha_G}{12(1-\alpha_F\alpha_G)}\right] \\
&\quad= \frac{(1-\alpha_F)(-\alpha_F^2\alpha_G^2 - \alpha_F\alpha_G^2 + 2\alpha_F\alpha_G - 2\alpha_F + 2\alpha_G - 10)}{12(1-\alpha_F\alpha_G)^2} \\
&\quad\leq \frac{(1-\alpha_F)(-\alpha_F^2\alpha_G^2 - \alpha_F\alpha_G^2 - 2\alpha_F - 6)}{12(1-\alpha_F\alpha_G)^2} \\
&\quad\leq 0,
\end{aligned}$$

so for any fixed α_F, the above ratio is minimized as $\alpha_G \to 1$. Hence, choosing α_F that minimizes

$$\lim_{\alpha_G \to 1} \frac{12 - \alpha_F^2 \alpha_G^2 + 2\alpha_F^2 \alpha_G - 4\alpha_F \alpha_G + \alpha_G^2 - 10\alpha_G}{12(1 - \alpha_F \alpha_G)} = \frac{3 + \alpha_F^2 - 4\alpha_F}{12(1 - \alpha_F)} = \frac{3 - \alpha_F}{12},$$

we take $\alpha_F \to 1$. Therefore, for sufficiently small δ and suitable choices of $0 < \alpha_F \leq \alpha_G < 1$, we have that

$$\frac{\text{SW}}{\text{FB-SW}} < \frac{1}{6} + \epsilon.$$

a

Appendix G: Tokenized Bandit for LLM Decoding and Alignment

G.1 Related work

Multi-armed bandit and variants. Dating back to [240], multi-armed bandits and its linear bandit variant have been central to modern recommender system [32, 178, 195]. There exists a few works investigating its variant to tree search model [176], analyzing efficient algorithms under certain structural assumption, *e.g.,* smoothness of tree.[1]

Decoding-time alignment. There are a line of works that use heuristic methods for decoding-time alignment using external modules that guide the decoding process. [168] suggested aligning the LLM in the decoding time to align the likelihood and utility based on some external information. [170] present a token level Markov decision process (MDP) framework called controlled decoding using a separate prefix scorer rule that is trained to learn a value function from rewards. [169] present a decoding scheme as a heuristic-guided search process to align various preferences. [171] propose a transfer decoding to transfer-learn the unknown Q function in the token level MDP, and analyze the upper bound on the gap between the optimal policy and their algorithm. Meanwhile, our decoding-time alignment problem is solely framed as a bandit problem not MDP.

Foundation of alignment. Recent works have provided some theoretical foundations on LLM alignment. [192] derived a closed-form characterization of the optimal KL-constrained reinforcement learning (RL) alignment problem, and conclude that the famous best-of-N policy and KL-constrained RL is asymptotically equivalent. [241] analyze the upper bound on the KL divergence of best-of-N pol-

[1] More detailed discussion can be found in Appendix G.3.

icy. We remark that our LLM alignment problem is formulated in a largely different manner (see Section 8.5.1), thus less relevant to these literature.

Foundation of decoding. [242] analyze the perplexity of several sampling methods assuming that the token probabilities follow a Zipf distribution. [174] provided a theoretical justification on using truncation to avoid the excessive probability assignment on unreliable tail, empirically observed by [190]. Recently, [175] propose a decoding problem as a two-player game, and provide near-optimal strategies that encompass greedy decoding and some of its variants. Our LLM decoding problem is formulated as a sequence function maximization problem to maximize the posterior probability given a user query, which is largely different from above.

G.2 k-DDMC and Lookahead Decoding

Interestingly, we can relax the DDMC assumption and obtain analogous results by replacing the greedy decoding part with the *lookahead* decoding. In k-lookahead decoding, it generates all possible combination of the following k tokens, and select the token sequence that maximizes the utility (probability). This allows us to recover the fictitious extension argument used in both the analysis of TLB and TMAB.

First, k-DDMC is defined as follows:

Assumption G.2.1 (k-DDMC)**.** *A sequence function $u(\cdot)$ is k-DDMC if*

$$u(x_t, \mathbf{y} : \boldsymbol{a}) - u(x_t, \mathbf{z} : \boldsymbol{a}) \leq u(x_t, \mathbf{y}) - u(x_t, \mathbf{z}),$$

for any two finite sequences $\mathbf{x}, \mathbf{y}$ with the same length $|\mathbf{x}| = |\mathbf{y}|$, any sequence $\boldsymbol{a} \in \mathcal{V}^*$ with $|\boldsymbol{a}| = k$,

and any query x_t such that LHS is nonnegative.

ALGORITHM 10: k-Lookahead-EOFUL

Input: Token set $\mathcal{V}$, linear contextual bandit algorithm ALG
Initialize $\mathbf{y} \leftarrow \emptyset$, $\hat{\theta}_1 \leftarrow \vec{0}$, $C_1 = \{\hat{\theta}\}$, $\mathbf{y} = \emptyset$
for $t = 1, 2, \ldots, T$ **do**
  User arrives and submits query x_t
  for $k = 1, 2, \ldots, L-1$ **do**
    Compute $\mathbf{z}^* = \operatorname{argmax}_{\mathbf{z} \in \mathcal{V}^*, |\mathbf{z}|=k, \theta \in C_t} \langle \theta, e(x_t, \mathbf{y} : \mathbf{z}) \rangle$
    Set $\mathbf{y} \leftarrow \mathbf{y} : \mathbf{z}^*$
    if $(\mathbf{z}^*)^{(-1)} = \texttt{EOS}$ **then** Break;
  Submit $\mathbf{y}$ and observe reward r_t
  Compute $\hat{\theta}_{t+1}$ by (8.1) and C_{t+1} by (8.2)

ALGORITHM 11: k-LookaheadETC

Input: Token set $\mathcal{V}$, exploration parameter N
Initialize $\mathbf{y} \leftarrow \emptyset$
for $k = 1, 2, \ldots, L$ **do**
  if $|\mathbf{y}| = L - 1$ **then**
    $\mathbf{y} \leftarrow \mathbf{y} : \texttt{EOS}$
    Break
  for $\mathbf{z} \in \mathcal{V}^*$ *with* $|\mathbf{z}| = k$ **do**
    Submit $\mathbf{y} : t : \texttt{EOS}$ for N times and observe rewards
    Set $\bar{r}(\mathbf{y} : \mathbf{z})$ be the averaged cumulative rewards
  Select $\mathbf{z}^* = \operatorname{argmax}_{\mathbf{z} \in \mathcal{V}^*, |\mathbf{z}|=k} \bar{r}(\mathbf{y} : \mathbf{z})$
  Set $\mathbf{y} \leftarrow \mathbf{y} : \mathbf{z}^*$
  if $(\mathbf{z}^*)^{(-1)} = \texttt{EOS}$ **then**
    Break
For the rest of the rounds, submit $\mathbf{y}$

Correspondingly, we can modify our algorithms as presented in Algorithm 10 and Algorithm 11.

Now we given an outline of how one could extend our analysis for EOFUL and GreedyETC. Notice that the key part for the regret analysis is to compare an algorithm's choice and optimal algorithm's choice $\mathbf{a}_t^{(1:i)}$ and $\mathbf{o}_t^{(1:i)}$, respectively, up to i-th token at round t by considering a fictitious extension $\mathbf{f} = \mathbf{a}_t^{(1:i)} : \mathbf{o}_t^{(i+1)}$. With k-DDMC, we extend this analysis by considering alternative fictitious extension such that $\mathbf{f} = \mathbf{a}_t^{(1:i)} : \mathbf{o}_t^{(i+1:i+k)}$. With such extension, until level $j \leq k$, our algorithm is guaranteed to decode the (empirically) optimal token sequence given the estimates.

Then, for level $j > k$ such that $j = i + k$, we have:

$$\begin{aligned}
\mathbf{o}_t^{(1:j)} - \mathbf{a}_t^{(1:j)} = \mathbf{o}_t^{(1:i+k)} - \mathbf{a}_t^{(1:i+k)} &= \mathbf{o}_t^{(1:i+k)} - \mathbf{f} + \mathbf{f} - \mathbf{a}_t^{(1:i+k)} \\
&\leq \mathbf{o}_t^{(1:i)} - \mathbf{a}_t^{(1:i)} + \mathbf{f} - \mathbf{a}_t^{(1:i+k)} && \text{(By } k\text{-DDMC)} \\
&\leq \mathbf{o}_t^{(1:i)} - \mathbf{a}_t^{(1:i)}, && \text{(By optimistic choice of algorithms)}
\end{aligned}$$

thereby connecting j-level regret to $j - k$-level regret. With telescoping arguments and following our analysis for EOFUL and GreedyETC, one can prove that k-Lookahead-EOFUL and k-LookaheadETC can obtain regret sublinear in T and linear in L. We leave the detailed analysis to the readers.

G.3 Connection to Bandit Tree Search problem

In bandit tree search (BTS) problem by [176], there exists a binary tree with depth d. Each leaf node i is assigned a random variable X_i with bounded support $[0, 1]$ and expectation μ_i. At each round $t \in [T]$, the DM chooses a leaf node I_t and observe the random reward X_{I_t}. The DM, however, is unknown about the tree structure and needs to explore it. Likewise, the DM irrevocably selects each child in the tree until it reaches a leaf node. A static and decision version of this problem is to decide whether there exists a leaf node l whose utility function $v(l) \geq A$ for some $A \geq 0$. Likewise, a static and decision version of TMAB is to decide whether there exists a subsequence $\mathbf{y}$ with $u(\mathbf{y}) \geq A$ for some $A \geq 0$ in polynomial time. We will see that static decision versions of these problems can be reduced from each other. Illustrative examples of such reductions are provided in Figure G.1.

The following theorem formally argues that one can reduce BTS to TMAB:

Theorem G.3.1. *An instance of static and decision version of BTS can be reduced to an instance of*

static and decision version of TMAB and monotone utility function.[2]

Proof. Let $\mathcal{L}$ be the set of leads and d be the depth of the tree in BTS. For each leaf node $l \in \mathcal{L}$, we construct a corresponding token $\tau(l) \in \mathcal{V}$ and add `EOS` token, *i.e.,* $\mathcal{V} = \{\tau(l) : l \in \mathcal{L}\} \cup \{\texttt{EOS}\}$. Let $v(\cdot)$ be the value function for BTS that maps leaf to nonnegative real values. We write $d(x)$ to denote the depth of the any (possibly internal) node x in the instance of BTS. Let A_l be the set of nodes in depth l in BTS. We arbitrarily define an injection σ_l from A_l to $\mathcal{V} \setminus \{\texttt{EOS}\}$. We say that $\tau \in \mathcal{V}$ is *covered* at level l if there exists a node x with depth l in BTS such that $\sigma_l(x) = \tau$.

Now, we construct sequence function $u : \mathcal{V}^* \to \mathbb{R}_{\geq 0}$ as follows: for every $\mathbf{y} \in \mathcal{V}^*$

1. If $\mathbf{y}^{(-1)} \neq 0$, then $u(\mathbf{y}) = 0$.
2. If $\mathbf{y}^{(-1)} \neq 0$ and if every $\mathbf{y}^{(k)}$ for $k \in [|\mathbf{y} - 1|]$ is covered, then $u(\mathbf{y}) = v(\sigma_l^{-1}(\mathbf{y}^{(-2)}))$.
3. If $\mathbf{y}^{(-i)} = \texttt{EOS}$ for $i = 1, 2, \ldots, k$ for $k \geq 2$ and $\mathbf{y}^{(-k)} \neq \texttt{EOS}$, then $u(\mathbf{y}) = u(\mathbf{y}^{(1:|\mathbf{y}|-(k-1))})$.

It is straightforward to check that No instance of the static and decision version of BTS enforces No instance to that of TMAB, since no sequence would have utility larger than or equal to A. Consider Yes instance of the static and decision version of BTS and let l be the leaf node with $v(l) \geq A$. Now we define $\mathbf{y}$ to satisfy:

1. $|\mathbf{y}| = d(l) + 1$.
2. $\mathbf{y}^{(k)} = \sigma_k(l)$ for every $k \in [d(l)]$.
3. $\mathbf{y}^{(d(l)+1)} = \texttt{EOS}$.

Then, due to our construction of the utility function above, we have $u(\mathbf{y}) = v(l) \geq A$, finishiing the reduction.

[2]The reduction can easily be extended to k-ary tree version of BTS.

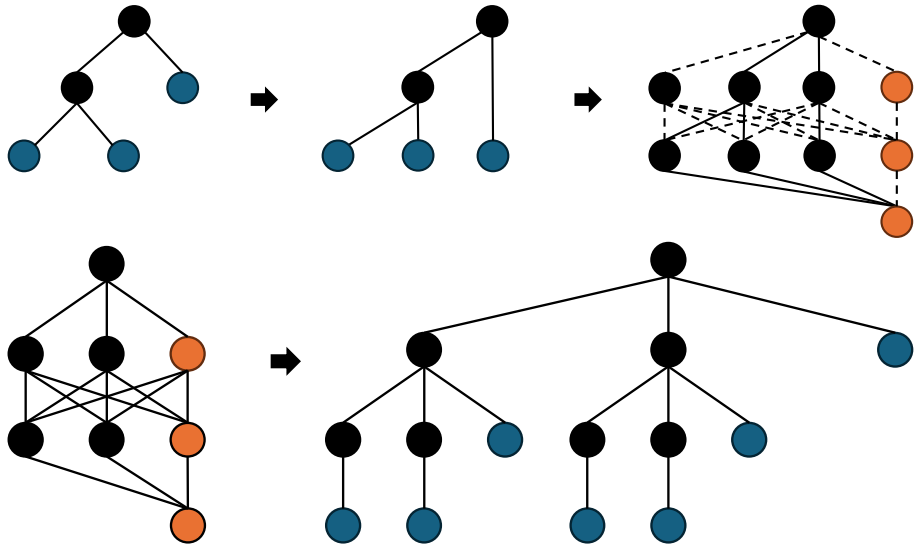


Figure G.1: Illustration of reductions. The above figure is an example of reduction from BTS to TMAB, and the below is from TMAB to generalized BTS (with k-ary tree). In both figures, blue nodes represent leaf nodes and orange node denotes `EOS` token. Dotted edge in the rightmost graph implies that the sequence contains such edge would have utility zero. Thus, only the sequences without dotted lines have (potentially) nonzero utility corresponding to the utility at the leaf node in the original instance of BTS. In the below figure, each path to `EOS` in the instance of TMAB becomes a leaf node in the instance of generalized BTS.

Now it remains to prove that the constructed utility function is monotonicity as defined in Assumption 8.2.1. This naturally follows by step 3 in the construction of $u(\cdot)$, since adding more `EOS` does not change the utility.

a

We remark that one might think that TMAB is not a search problem unlike BTS since we know the set of feasible sequences in advance, but it can indeed be viewed as a search problem since the ultimate goal is to find a node with larger, nonzero, utility.

Theorem G.3.2. *An instance of static and decision version of TMAB can be reduced to an instance of static and decision version of TS.*

Proof. We sort the token set $\mathcal{V}$ and index by $\sigma : \mathcal{V} \to \{1, 2, \ldots, n\}$ where such that $\sigma(\texttt{EOS}) = n$.

Given an instance of BTS problem with $\mathcal{V}$ we construct n-ary tree. For depth d of the tree, there will be $n-1$ NORMAL nodes who have n children, and one END node with no children. We will arbitrarily assign index $i \in [n-1]$ to each normal node and n to the END node.

Consider any complete sequence $\mathbf{y}$ with length l. Then, we can select a corresponding leaf node l such that along the path p to l, the node in depth d in path p has index $\sigma(\mathbf{y}^{(d)})$. Since $\mathbf{y}$ is complete, the corresponding leaf node l is indeed a leaf since its last index is $\sigma(\mathbf{y}^{(-1)}) = n$, which is END node that does not have any children. Thus, we assign utility $v(l) = u(\mathbf{y})$. Then, the reduction easily follows, and we leave the details to the readers. a

Discussion G.3.3. Accordingly, one natural question is what it means by the binary tree in BTS having DDMC structure. As noted in the proof and Figure G.1, the DDMC assumption in TS means that w

DDMC assumption in BTS, smoothness in TMAB. Given the close connection between two problems, one might wonder what the DDMC assumption means in BTS. If we consider k-ary tree version of BTS, then we can imagine an ordering between children for each node, let the index be $1, 2, \ldots, n$. Then, in BTS, our DDMC assumption asserts that each two nodes at the same depth append the children with the same index i, the resulting utility gap only decreases compared to the utility gap between the parent nodes.

On the other hand, [176] presents a rather immediate and strong assumption on the tree, named smoothness, stated as follows:

Assumption G.3.4 (Smoothness)**.** *Given an instance of BTS, for any depth $d < L$, there exists $\delta_d > 0$ such that for at least one optimal node i of depth d (that leads to the optimal leaf node), for all leaves*

j in the subtree of i, it follows that

$$u(l^*) - u(u_j) \leq \delta_d,$$

where l^* denotes the optimal leaf node.

That is, it upper bounds the utility gap at depth d by the parameter δ_d, and provide a regret upper bound parameterized by such δ_d. This could analogously be applied in our TMAB problem, by upper bounding the utility gap between any sequence of length l and the subsequence of length l that leads to the optimal subsequence by some parameters δ_l. On the other hand, our DDMC assumption does not introduce any such upper bound on the utility gap directly, but rather imposes a minor structural relation between sequences with length difference 1. Thus, we believe that our DDMC assumption can also play a crucial role in the BTS problem, given that it imposes certain structures on the tree implicitly, and the empirical evidence captured in Section 8.5.

G.4 Proofs

G.4.1 Proof of Proposition 8.3.3

Proof of Proposition 8.3.3. Let $\mathcal{L}$ be the set of sequences with length L so that $|\mathcal{L}| = 2^L$ and we arbitrarily index the elements in $\mathcal{L}$ and write $\mathcal{L} = \{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_{2^L}\}$. Let o be the one-hot vector who has one only at the first coordinate and otherwise zero, and set $\theta = o$. Consider a choice of the query sequence $(x_t)_{t\in[T]}$ and the embedding function such that $e(x_t, \mathbf{y})$ has a unique subsequence $\mathbf{y} \in \mathcal{L}$ such that $e(x_t, \mathbf{y}) = o$, and otherwise $e(x_t, \mathbf{y}) = 0$ for every other $\mathbf{y} \in \mathcal{V}^*$. Further, assume that the sequence $(x_t)_{t\in[T]}$ ensures that such unique $\mathbf{y} \in \mathcal{L}$ is selected uniformly at random over

$\mathcal{L}$. Let the chosen unique such vector at round t be $\mathbf{y}_t$. Note that this is indeed possible by having large enough candidates for x_t whose cardinality scales as $\Theta(T)$ and for each x_t it independently and uniform randomly selects the unique such $\mathbf{y} \in \mathcal{L}$ with $e(x_t, \mathbf{y}) = 0$. It is straightforward to check that the corresponding utility function satisfies Assumption 8.2.1. Observe that such instance is in fact a randomization over deterministic instances each of which deterministically choose $\mathbf{y}_t \in \mathcal{L}$ for each $t \in [T]$. Thus, by Yao's minimax principle [243], it suffices to consider a class of deterministic algorithms.

In this instance, there always exists a sequence $\mathbf{y} \in \mathcal{L}$ such that $e(x_t, \mathbf{y}) = o$, and thus the optimal decoding reward in hindsight is to always select $\mathbf{y} \in \mathcal{L}$ with $e(x_t, \mathbf{y}) = o$, which induces the cumulative reward of $T \cdot \langle o, o \rangle = T$. On the other hand, since the choice of $\mathbf{y}_t$ is completely independent from the previously chosen sequences and realization of random variables therein, any deterministic algorithm's probability to choose $\mathbf{y}_t$ is at most the random guess over 2^{L-1} elements, resulting in the cumulative reward of $T \cdot \langle o, o \rangle / 2^{L-1} = T/2^{L-1}$. Note here that this is 2^{L-1} instead of 2^L since once the algorithm selects $\mathbf{y}_t^{(1:L-1)}$, it reward is upper bounded by an algorithm that deterministically chooses $\mathbf{y}_t^{(i)}$. Thus, the regret is at least $T(1 - 1/2^{L-1})$. a

G.4.2 Proof of Theorem 8.3.9

We here provide the proof of Theorem 8.3.9. Statement and proofs of the required technical lemmas are deferred to the end of this section.

Proof of Theorem 8.3.9. Throughout, we assume that `Clean` holds defined in Lemma G.4.1, which occurs with probability at least δ.

Fix a round t, and let $\tilde{\theta}_t^{(k)} \in C_t$ be the optimistic estimator chosen at round t for k-th token.

Let $\mathbf{o}_t$ be the optimal sequence chosen by the optimal algorithm in hindsight at round t. For ease of exposition, in representing utility given sequence $\mathbf{y} \in \mathcal{V}^*$ and query x_t, we abuse $\langle \theta, \mathbf{y} \rangle$ to denote $\langle \theta, e(x_t, \mathbf{y}) \rangle$ unless confusion arises, *i.e.,* consider $\mathbf{y}$ to be the resulting embedded vector $e(x_t, \mathbf{y}) \in \mathbb{R}^d$.

(Length equalization) We will first define a level-k regret that computes the difference of utility between the k-prefix of $\mathbf{o}_t$ and $\mathbf{y}_t$. However, this is not well-defined for every $k \in \max(|\mathbf{o}_t|, |\mathbf{y}_t|)$ if $\mathbf{o}_t$ and $\mathbf{y}_t$ have different lengths. Thus, we consider the *length equalization* operation on $\mathbf{o}_t$ and $\mathbf{y}_t$ that constructs extensions $\tilde{\mathbf{o}}$ and $\tilde{\mathbf{a}}$ of $\mathbf{o}$ and $\mathbf{a}$ respectively to have the same length in what follows.

We have two cases: (i) if $|\mathbf{a}| > |\mathbf{o}|$, then we append multiple `EOS` tokens at the end of $\mathbf{o}$ and construct $\tilde{\mathbf{o}}$ so that $|\tilde{\mathbf{o}}| = |\mathbf{a}|$, and (ii) otherwise $|\mathbf{a}| < |\mathbf{o}|$, then we do the same operation for $\mathbf{a}$ to construct $\tilde{\mathbf{a}}$ with $|\tilde{\mathbf{a}}| = |\tilde{\mathbf{o}}|$. Then, we extend both sequences $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{o}}$ to exactly have the length L by appending `EOS`. By Assumption 8.2.1, we have $u(\mathbf{o}) = u(\tilde{\mathbf{o}})$ and $u(\mathbf{a}) = u(\tilde{\mathbf{a}})$. Thus, we can safely deal with the extended sequences with the same length as the regret remains the same. For ease of exposition, we override $\mathbf{a}_t \leftarrow \tilde{\mathbf{a}}_t$ and $\mathbf{o}_t \leftarrow \tilde{\mathbf{o}}_t$ for every $t \in [T]$.

(Level-k regret and fictitious extension) Now, we can formally define level-k regret as follows:

$$\textsc{Reg}_t^{(k)} = \left\langle \theta, \mathbf{o}_t^{(1:k)} \right\rangle - \left\langle \theta, \mathbf{y}_t^{(1:k)} \right\rangle.$$

Now we observe the connection between the level-k regret with different k. For $k = 2, \ldots, L$, we consider an one-step *fictitious extension* of $\mathbf{y}_t^{(1:k-1)}$ that appends the optimal sequence $\mathbf{o}_t$'s k-th token to $\mathbf{y}_t$'s (k-1)-th prefix, *i.e.,*

$$\mathbf{f}_t^{(1:k)} = \mathbf{y}_t^{(1:k-1)} : \mathbf{o}_t^{(k)}$$

Notice that $\mathbf{o}_t^{(1:k)}$ and $\mathbf{f}_t^{(1:k)}$ shares the same common suffix $\mathbf{o}^{(k)}$. Thus, by DDMC assumption, behold:

$$\begin{aligned}\left\langle \theta, \mathbf{o}_t^{(1:k)} \right\rangle - \left\langle \theta, \mathbf{f}_t^{(1:k)} \right\rangle &\le \left\langle \theta, \mathbf{o}_t^{(1:k-1)} \right\rangle - \left\langle \theta, \mathbf{f}_t^{(1:k-1)} \right\rangle \\ &= \left\langle \theta, \mathbf{o}_t^{(1:k-1)} \right\rangle - \left\langle \theta, \mathbf{y}_t^{(1:k-1)} \right\rangle \qquad \text{(G.1)} \\ &= \textsc{Reg}_t^{(k-1)} \qquad \text{(G.2)}\end{aligned}$$

(Regret decomposition) Now, for $k \ge 2$, the level-k regret can be decomposed as follows:

$$\begin{aligned}\textsc{Reg}_t^{(k)} &= \left\langle \theta, \mathbf{o}_t^{(1:k)} \right\rangle - \left\langle \theta, \mathbf{y}_t^{(1:k)} \right\rangle \\ &= (a) + (b) + (c),\end{aligned}$$

where

$$\begin{aligned}(a) &= \left\langle \theta, \mathbf{o}_t^{(1:k)} \right\rangle - \left\langle \theta, \mathbf{f}_t^{(1:k)} \right\rangle \\ (b) &= \left\langle \theta, \mathbf{f}^{(1:k)} \right\rangle - \left\langle \tilde{\theta}_t^{(k)}, \mathbf{y}_t^{(1:k)} \right\rangle \\ (c) &= \left\langle \tilde{\theta}_t^{(k)}, \mathbf{y}_t^{(1:k)} \right\rangle - \left\langle \theta, \mathbf{y}_t^{(1:k)} \right\rangle\end{aligned}$$

By (G.2), we have $(a) \le \textsc{Reg}_t^{(k-1)}$.

Since our algorithm chooses $\mathbf{y}_t^{(k)}$ instead of $\mathbf{o}_t^{(k)}$ given $\mathbf{y}_t^{(1:k-1)}$ for k-th token, by the optimistic choice we have $(b) \le 0$.[3]

Let us denote the term in (c) by $W_t^{(k)}$ for $k \in [M]$.

[3] We note that this argument remains the same regardless of the length equalization process above. For instance, even if $\mathbf{y}_t^{(k-1)} = \texttt{EOS}$ and thus $\mathbf{y}_t^{(k)} = \texttt{EOS}$, by Assumption 8.2.1, appending `EOS` is the greedy choice.

Thus, we have

$$\textsc{Reg}_t^{(k)} \le \textsc{Reg}_t^{(k-1)} + W_t^{(k)}.$$

Telescoping over k:

$$\textsc{Reg}_t \le \sum_{k=2}^{L} W_t^{(k)} + \textsc{Reg}_t^{(1)}.$$

Define $w_t(\mathbf{x}) = \sqrt{\mathbf{x}^\top \Sigma_t^{-1} \mathbf{x}}$ for $\mathbf{x} \in \mathbb{R}^d$. Analogous to the standard regret analysis, we eventually use the sum of square regret arguments.

Now we invoke Lemma G.4.3 to obtain the following two useful upper bounds for $\textsc{Reg}_t^{(1)}$ and $W_t^{(k)}$. First for $\textsc{Reg}_t^{(1)}$:

$$\begin{aligned}
\textsc{Reg}_t^{(1)} &= \left\langle \theta, \mathbf{o}^{(1)} \right\rangle - \left\langle \theta, \mathbf{y}^{(1)} \right\rangle \\
&\le \left\langle \tilde{\theta}_t^{(1)}, \mathbf{y}^{(1)} \right\rangle - \left\langle \theta, \mathbf{y}^{(1)} \right\rangle \\
&= \left\langle \tilde{\theta}_t^{(1)}, \mathbf{y}^{(1)} \right\rangle - \left\langle \hat{\theta}_t, \mathbf{y}^{(1)} \right\rangle + \left\langle \hat{\theta}_t, \mathbf{y}^{(1)} \right\rangle - \left\langle \theta, \mathbf{y}^{(1)} \right\rangle \\
&\le 2\sqrt{\beta_t} w_t(\mathbf{y}^{(1)}) \\
&\le 2\sqrt{\beta_t} \min(w_t(\mathbf{y}^{(1)}), 1),
\end{aligned}$$

where the first inequality follows from the optimistic choice of our algorithm, and the second inequality follows from Lemma G.4.3 since both $\tilde{\theta}_t^{(1)}, \theta \in C_t$ under `Clean`. The last inequality holds since the regret is at most 2 for any round.

Next, similarly for $W_t^{(k)}$ for $k = 2, 3, \ldots, L$:

$$\begin{aligned} W_t^{(k)} &= \left\langle \tilde{\theta}_t^{(k)}, \mathbf{y}_t^{(1:k)} \right\rangle - \left\langle \theta, \mathbf{y}_t^{(1:k)} \right\rangle \\ &= \left\langle \tilde{\theta}_t^{(k)}, \mathbf{y}_t^{(1:k)} \right\rangle - \left\langle \hat{\theta}_t, \mathbf{y}_t^{(1:k)} \right\rangle + \left\langle \hat{\theta}_t, \mathbf{y}_t^{(1:k)} \right\rangle - \left\langle \theta, \mathbf{y}_t^{(1:k)} \right\rangle \\ &\leq 2\sqrt{\beta_t} \min(w(\mathbf{y}_t^{(1:k)}), 1). \end{aligned}$$

Combining two, we have

$$\textsc{Reg}_t \leq 2\sqrt{\beta_t} \left(\min(w(\mathbf{y}_t), 1) + \sum_{k=1}^{L-1} \min(w(\mathbf{y}_t^{(1:k)}), 1) \right).$$

Now, by Assumption 8.3.6, we have

$$w(\mathbf{y}_t^{(1:k)}) = \sqrt{(\mathbf{y}_t^{(1:k)})^\top \Sigma_t^{-1} \mathbf{y}_t^{(1:k)}} \leq c\sqrt{\mathbf{y}_t^\top \Sigma_t^{-1} \mathbf{y}_t} = w(\mathbf{y}_t).$$

Hence, we have

$$\textsc{Reg}_t \leq 2\sqrt{\beta_t}(1 + cL) \min(w_t(\mathbf{y}_t), 1).$$

(Sum of squares regret) Squaring:

$$\textsc{Reg}_t^2 \leq 4\beta_t(1 + cL)^2 \min(w_t(\mathbf{y}_t)^2, 1),$$

and

$$\begin{aligned}\sum_{t=1}^{T}\textsc{Reg}_t^2 &\le \sum_{t=1}^{T}4\beta_t(1+cL)^2\min(w_t(\mathbf{y}_t)^2,1)\\ &\le 4\beta_T(1+cL)^2\sum_{t=1}^{T}\min(w_t(\mathbf{y}_t)^2,1)\\ &\le 4\beta_T(1+cL)^2\sum_{t=1}^{T}\ln(1+w_t(\mathbf{y}_t)^2),\end{aligned}$$

where the last inequality uses the fact that for any $x\in[0,1]$, $\ln(1+x)\ge x/(1+x)\ge x/2$, thus when $w_t(\mathbf{y}_t)^2\le 1$, we have $w_t(\mathbf{y}_t)^2\le 2\log(1+w_t(\mathbf{y}_t)^2)$ and otherwise if $w_t(\mathbf{y}_t)^2>1$, we have

$$4\beta_t=\frac{4}{\log 2}\beta_t\log 2\le\frac{4}{\log 2}\beta_t\log(1+w_t(\mathbf{y}_t)^2).$$

Then, we have

$$\begin{aligned}\sum_{t=1}^{T}\textsc{Reg}_t^2 &\le 4\beta_T(1+cL)^2\sum_{t=1}^{T}\ln(1+w(\mathbf{y}_t)^2)\\ &\le 4\beta_T(1+cL)^2 d\ln\left(1+\frac{T}{d\lambda}\right) && \text{(By Lemma G.4.4 and G.4.5)}\end{aligned}$$

Finally, by Cauchy-Schwartz inequality,

$$\begin{aligned}\sum_{t=1}^{T}\textsc{Reg}_t &\le \sqrt{T\cdot\sum_{t=1}^{T}\textsc{Reg}_t^2}\\ &\le \sqrt{4T(1+cL)^2\beta_T d\ln\left(1+\frac{T}{d\lambda}\right)}\\ &= O(cL\sqrt{dT\ln T}), && \text{(Since } \beta_T=O(\lambda))\end{aligned}$$

a

Now we provide the statements of lemmata. We begin with the standard concentration argument stating that $\theta \in C_t$ with high probability:

Lemma G.4.1 (Confidence bound)**.** *Define clean event:*

$$\texttt{Clean} = \{\theta \in C_t, \forall t \in [T]\}.$$

Then, $\Pr[\texttt{Clean}] \geq 1 - \delta$.

Lemma G.4.2 (Self-normalized bound for martingale [178])**.** *Let* $\{F_t\}_{t=0}^{\infty}$ *be a filtration. Let* $\{\eta_t\}_{t=1}^{\infty}$ *be a real-valued stochastic process such that* η_t*is* F_t*-measurable and* η_t *is conditionally* R*-sub-Gaussian for some* $R \geq 0$*. Let* $\{X_t\}_{t=1}^{\infty}$ *be an* $\mathbb{R}^d$*-valued stochastic process such that* X_t *is* F_{t-1}*-measurable. Assume that* V *is a* $d \times d$ *positive-definite matrix. For any* $t \geq 0$*, define:*

$$\bar{V}_t = V + \sum_{s=1}^{t} X_s X_s^\top, \; S_t = \sum_{s=1}^{t} \eta_s X_s.$$

Then for any $\delta > 0$*, with probability at least* $1 - \delta$*, for all* $t \geq 0$*,*

$$\|S_t\|^2_{\bar{V}_t^{-1}} \leq 2R^2 \log\left(\frac{\det(\bar{V}_t^{1/2} \det(V)^{1/2})}{\delta}\right).$$

For $\mathbf{z} \in \mathbb{R}^d$, define $w_t(\mathbf{z}) = \sqrt{\mathbf{z}^\top \Sigma_t \mathbf{z}}$. We then need the following several lemmata to proceed the sum of squares regret bound:

Lemma G.4.3. *For any* $\mathbf{z} \in \mathbb{R}^d$ *and* $\vartheta \in C_t$*, we have*

$$\left\langle \vartheta - \hat{\theta}_t, \mathbf{z} \right\rangle \leq 2\sqrt{\beta_t} w_t(\mathbf{z}) \tag{G.3}$$

Lemma G.4.4. *The following equation holds:*

$$\det(\Sigma_T) = \det(\Sigma_1) \prod_{t=1}^{T} (1 + w_t(\mathbf{y}_t)^2).$$

Lemma G.4.5. *We have:*

$$\log \frac{\det(\Sigma_T)}{\det(\Sigma_1)} = d \log \left(1 + \frac{T}{d\lambda}\right).$$

The proofs are mostly similar to the standard linear bandit analysis [178], but is tailored to our setting. Thus, for completeness, we provide the proofs of some of the lemmata:

Proof of Lemma G.4.1. Note that the reward at round t is $r_t = \langle \theta, e(x, \mathbf{y}_t) \rangle + \eta_t$. For concise exposition, we abuse the notation and write $\mathbf{y}_t = e(x, \mathbf{y}_t)$.

$$\begin{aligned}
\hat{\theta}_t - \theta &= \Sigma_t^{-1} \sum_{i=1}^{t} r_i \mathbf{y}_i - \theta \\
&= \Sigma_t^{-1} \sum_{i=1}^{t} \mathbf{y}_i (\langle \theta, \mathbf{y}_i \rangle + \eta_t) - \theta \\
&= \Sigma_t^{-1} \left(\mathbf{y}_i \mathbf{y}_i^\top\right) \theta - \theta + \Sigma_t^{-1} \sum_{i=1}^{t} \eta_i \mathbf{y}_i \\
&= -\lambda \Sigma_t^{-1} \theta + \Sigma_t^{-1} \sum_{i=1}^{t} \eta_i \mathbf{y}_i.
\end{aligned}$$

Notice that

$$\sqrt{(\hat{\theta}_t - \theta)\Sigma_t(\hat{\theta}_t - \theta)^\top} = \left\|\Sigma_t^{1/2}(\hat{\theta}_t - \theta)\right\|$$

Expanding:

$$\begin{aligned}\left\|\Sigma_t^{1/2}(\hat{\theta}_t - \theta)\right\| &\leq \left\|\lambda\Sigma_t^{-1/2}\theta\right\| + \left\|\Sigma_t^{-1/2}\sum_{i=1}^{t}\eta_i\mathbf{y}_i\right\| \\ &\leq \sqrt{\lambda}\,\|\theta\| + \sqrt{2\log\frac{\det(\Sigma_t)\det(\Sigma_1)^{-1}}{\delta_t}},\end{aligned}$$

where the last inequality uses $\Sigma_t^{-1} \leq 1/\lambda$ due to the construction.

Note first that $\hat{\theta}_1 = \vec{0} \in C_1$ surely. For $t \geq 2$, we set $\delta_t = 3/(\pi^2 t^2) \cdot 2\delta$. Then, using union bound, one can deduce that

$$\Pr\left[\,\theta \in C_t, \forall t \in [T]\,\right] \geq 1 - \delta.$$

a

Proof of Lemma G.4.3. By Cauchy-Schwarz inequality and the fact that $\vartheta \in C_t$, we obtain the follow-

ing series of inequality:

$$\begin{aligned}
\left| (\vartheta - \hat{\theta}_t)^\top z \right| &= \left| (\vartheta - \hat{\theta}_t)^\top \Sigma_t^{1/2} \Sigma_t^{-1/2} z \right| \\
&= \left| (\Sigma_t^{1/2}(\vartheta - \hat{\theta}_t))^\top \Sigma_t^{-1/2} z \right| \\
&\le \left\| \Sigma_t^{1/2}(\vartheta - \hat{\theta}_t) \right\| \left\| \Sigma_t^{-1/2} z \right\| && \text{(Cauchy Schwarz inequality)} \\
&= \left\| \Sigma_t^{1/2}(\vartheta - \hat{\theta}_t) \right\| \sqrt{z^\top \Sigma_t^{-1} z} \\
&\le \sqrt{\beta_t z^\top \Sigma_t^{-1} z}. && \text{(Since } \vartheta \in C_t\text{)}
\end{aligned}$$

a

Proof of Lemma G.4.4. By construction of Σ_t, we have

$$\begin{aligned}
\det(\Sigma_{t+1}) &= \det(\Sigma_t + \mathbf{y}_t \mathbf{y}_t^\top) \\
&= \det\left(\Sigma_t^{1/2}\left(I + \Sigma_t^{-1/2}\mathbf{y}_t\mathbf{y}_t^\top\right)\Sigma_t^{1/2}\right) \\
&= \det(\Sigma_t)\det\left(I + \Sigma_t^{-1/2}\mathbf{y}_t\left(\Sigma_t^{-1/2}\mathbf{y}_t\right)^\top\right) \\
&= \det(\Sigma_t)\det\left(I + v_t v_t^\top\right),
\end{aligned}$$

where we write $v_t = \Sigma_t^{-1/2}\mathbf{y}_t$. Observe that $v_t^\top v_t = w_t(\mathbf{y}_t)^2$. Further,

$$(I + v_t v_t^\top)v_t = v_t + v_t(v_t^\top v_t) = (1 + w_t(\mathbf{z}_t)^2)v_t,$$

implying that $(1 + w_t(\mathbf{y}_t)^2)$ is an eigenvalue of $I + v_t v_t^\top$. Since $v_t v_t^\top$ is a rank one matrix, all other

eigenvalues of $I + v_t v_t^\top$ are one. Thus,

$$\det(\Sigma_{t+1}) = \det(\Sigma_t)(1 + w_t(\mathbf{y}_t)^2).$$

Telescoping:

$$\det(\Sigma_T) = \det(\Sigma_1) \prod_{t=1}^{T} (1 + w_t(\mathbf{y}_t)^2)$$

a

Proof of Lemma G.4.5. Let the eigenvalues of $\sum_{t=1}^{T} \mathbf{y}_t \mathbf{y}_t^\top$ be $\sigma_1, \sigma_2, \ldots, \sigma_d$. Note that

$$\sum_{i=1}^{k} \sigma_i = \text{Trace} \sum_{t=1}^{T} \mathbf{y}_t \mathbf{y}_t^\top = \sum_{t=1}^{T} \|\mathbf{y}_t\|^2 \leq T.$$

Thus, we have

$$\begin{aligned}
\log \det \frac{\Sigma_T}{\Sigma_1} &= \log \det \left(1 + \frac{1}{\lambda} \sum_{t=1}^{T} \mathbf{y}_t \mathbf{y}_t^\top \right) \\
&= \log \left(\prod_{i=1}^{d} \left(1 + \frac{\sigma_i}{\lambda} \right) \right) \\
&= d \log \left(\prod_{i=1}^{d} \left(1 + \frac{1}{\lambda} \right) \right)^{1/d} \\
&\leq d \log \left(\frac{1}{d} \sum_{i=1}^{d} (1 + \sigma_i / \lambda) \right) \\
&\leq d \log \left(1 + \frac{T}{d\lambda} \right),
\end{aligned}$$

which finishes the proof of the lemma. a

G.4.3 Proof of Theorem 8.5.1

Proof of Theorem 8.5.1. Let $\mathbf{o}$ be the optimal sequence and $\mathbf{g}$ be the resulting sequence from greedy decoding. Suppose not, *i.e.,* $\mathbf{o} \neq \mathbf{g}$ for sake of the contradiction, and let $\mathbf{g}$ be the resulting sequence by the greedy algorithm. Similar to the length equalization process in the proof of Theorem 8.3.9, we can extend $\mathbf{g}$ and $\mathbf{o}$ to have the same maximal length L. Now, it suffices to prove that for any $l \in [L]$, we have $u(\mathbf{o}^{(1:l)}) - u(\mathbf{g}^{(1:l)}) \leq 0$.

We proceed by induction. For $l = 1$, this obviously holds since greedy decoding selects the one that maximizes $u(\cdot)$. Assume this holds for some $l \in [L-1]$. Consider another sequence $\mathbf{f}$ that extends $\mathbf{g}^{(1:l)}$ by appending $\mathbf{o}^{(l+1)}$, *i.e.,* $\mathbf{f} = \mathbf{g}^{(1:l)} : \mathbf{o}^{(l+1)}$. Due to the greedy algorithm's nature, we have that

$$u(\mathbf{f}) - u(\mathbf{g}^{(1:l+1)}) \leq 0.$$

Thus, we have

$$u(\mathbf{o}^{(1:l+1)}) - u(\mathbf{g}^{(1:l+1)}) \leq u(\mathbf{o}^{(1:l+1)}) - u(\mathbf{f}).$$

Since $\mathbf{o}^{(1:l+1)}$ and $\mathbf{f}$ shares the same common at the end, by DDMC, we have

$$\begin{aligned} u(\mathbf{o}^{(1:l+1)}) - u(\mathbf{g}^{(1:l+1)}) &\leq u(\mathbf{o}^{(1:l+1)}) - u(\mathbf{f}) \\ &\leq u(\mathbf{o}^{(1:l)}) - u(\mathbf{g}^{(1:l)}) \\ &\leq 0, \end{aligned}$$

where the last inequality follows from the induction assumption. Thus, the greedy decoding is optimal.

a

G.4.4 Proof of Proposition 8.4.1

Proof of Proposition 8.4.1. It is straightforward to see that $m = |\mathcal{V}_L^*| = \Omega(n^L)$ since it has (roughly) at least $(n-1)^{L-1}$ sequences that ends with `EOS` and having depth L. Consider an instance of the stochastic m-armed bandit problem where each arm i is associated with a distribution F_i. Let us (arbitrarily) index each subsequence in $\mathcal{V}_L^*$ by $1, 2, \ldots, m$ and let $i(\mathbf{y})$ be index of $\mathbf{y}$. We construct an instance of TMAB such that each random reward $r_t(\mathbf{y})$ is an independent sample from $F_{i(\mathbf{y})}$. That is, whenever the DM submits a complete sequence $\mathbf{y} \in \mathcal{V}^*$, its random reward is sampled from $F_{i(\mathbf{y})}$ independently. Then, we have that $u(\mathbf{y}) = \int x dF_{i(\mathbf{y})}(x)$. We further assume that if we add more `EOS` to a complete sequence $\mathbf{y}$, the utility does not change, *i.e.,* have the same distribution. One can easily check that the utillity function satisfies Assumption 8.2.1. Then, in order to learn a reward for $\mathbf{y}$, any algorithm needs to run a bandit algorithm treating each $\mathbf{y}$ as a single arm. The standard information theoretic argument on regret lower bound gives us $\Omega(K) = \Omega(n^L)$.

a

G.4.5 Proof of Theorem 8.4.2

Proof of Theorem 8.4.2. Overall, we will prove that at any depth k during the execution algorithm, it will select the optimal subsequent token for depth $k+1$ with high probability. Let us begin with depth 1 for reader's convenience. Note that $\bar{r}(\mathbf{y} : \tau : \texttt{EOS})$ denotes the average of the cumulative reward obtained from exploring $\mathbf{y} : \tau : \texttt{EOS}$ for N rounds during the execution of the algorithm. For any

$\mathbf{y} \in \mathcal{V}^*$, define the clean event as follows:

$$\texttt{Clean}_{\mathbf{y}} = \{|\bar{r}(\mathbf{y}) - u(\mathbf{y})| \leq \sqrt{2\log T/N}\}.$$

By Hoeffding inequality, we have

$$\Pr\left[\,\texttt{Clean}_{\mathbf{y}}\,\right] \geq 1 - 2/T^4,$$

for any $\mathbf{y} \in \mathcal{V}^*$. Suppose our algorithm chooses token sequence $\mathbf{a}$ whereas the optimal algorithm chooses $\mathbf{o}$. If $\mathbf{o} = \texttt{EOS}$, then the problem boils down to a single-level scenario and the standard analysis of ETC algorithm carries over. Assume $\mathbf{o} \neq \texttt{EOS}$, *i.e.,* $\mathbf{o}$ considers at least one non-EOS token.

Due to our algorithm's choice, $\bar{r}(\mathbf{a}^{(1)} : \texttt{EOS}) \geq \bar{r}(\mathbf{o}^{(1)} : \texttt{EOS})$. Define

$$\texttt{Clean}^{(1)} = \cap_{\tau \in \mathcal{V}} \texttt{Clean}_{\emptyset:\tau:\texttt{EOS}},$$

i.e., the clean event over every token at depth 1. Taking union bound over $\mathbf{y} = \emptyset$ and $\tau \in \mathcal{V}$, we have

$$\Pr\left[\,\texttt{Clean}^{(1)}\,\right] \geq 1 - 2/T^3.$$

Under this event, we know that

$$\begin{aligned} u(\mathbf{a}^{(1)} : \texttt{EOS}) + \sqrt{2\log T/N} &\geq \bar{r}(\mathbf{a}^{(1)} : \texttt{EOS}) \\ &\geq \bar{r}(\mathbf{o}^{(1)} : \texttt{EOS}) \\ &\geq u(\mathbf{o}^{(1)} : \texttt{EOS}) - \sqrt{c\log T/N}. \end{aligned}$$

Thus, with probability $1 - 2/T^3$, we obtain $u(\mathbf{o}^{(1)}) - u(\mathbf{a}^{(1)}) \leq 2c \log T/N$. Define $d := \sqrt{2 \log T/N}$.

Now we will generalize this argument to prove that we have $u(\mathbf{o}^{(1:l)} : \texttt{EOS}) - u(\mathbf{a}^{(1:l)} : \texttt{EOS}) \leq d$ under an appropriate clean event occurring with high probability. Note that this is sufficient since $u(\mathbf{o}^{(1:l)} : \texttt{EOS}) = u(\mathbf{o}^{(1:l)})$ once $\mathbf{o}^{(1:l)}$ reaches EOS, and analogously for $\mathbf{a}$.

Subtly, we need to handle the case where $\mathbf{o}$ and $\mathbf{a}$ have different lengths. Similar to previous analysis, we equalize the length by adding more EOS tokens so that both $\mathbf{o}$ and $\mathbf{a}$ has length L. Then, we can safely deal with the extended sequences with the same length. We prove by induction on the length l of the subsequence of $\mathbf{o}$ and $\mathbf{a}$ by comparing $\mathbf{a}^{(1:l)}$ with $\mathbf{o}^{(1:l)}$. For every $i \in [L]$ and $i \geq 2$, we recursively define the following events:

$$\texttt{Clean}^{(1:i)} = \left(\cap_{j \in [i-1]} \texttt{Clean}^{(1:i-1)}\right) \cap \left(\cap_{\tau \in \mathcal{V}} \left(\texttt{Clean}_{\mathbf{a}^{(1:i-1)}:\tau:\texttt{EOS}}\right)\right) \cap \texttt{Clean}_{\mathbf{o}^{(1:i-1)}:\texttt{EOS}},$$

and define $\texttt{Clean}^{(1)}$. By abusing $\texttt{Clean}^{(1:0)} = \emptyset$ and $\mathbf{o}^{(1:0)} = \mathbf{a}^{(1:0)} = \emptyset$, note that we have $\texttt{Clean}^{(1:1)} = \texttt{Clean}^{(1)}$.

Now, assume for the sake of induction that $u(\mathbf{o}^{(1:i)} : \texttt{EOS}) - u(\mathbf{a}^{(1:i)} : \texttt{EOS}) \leq 2d$ holds under the clean event $\texttt{Clean}^{(1:i)}$. Now we will prove that under $\texttt{Clean}^{(1:i+1)}$, we have $u(\mathbf{o}^{(1:i+1)} : \texttt{EOS}) - u(\mathbf{a}^{(1:i+1)} : \texttt{EOS}) \leq 2(i+1)d$.

Similar to the analysis of EOFUL, we consider a fictitious extension of $\mathbf{f} = \mathbf{a}^{(1:i)} : \mathbf{o}^{(i+1)}$.

Due to the construction of clean event and our algorithm's optimistic choice, we have

$$\begin{aligned} u(\mathbf{a}^{(1:i+1)} : \texttt{EOS}) + \sqrt{c \log T/N} &\geq \bar{r}(\mathbf{a}^{(1:i+1)} : \texttt{EOS}) \\ &\geq \bar{r}(\mathbf{f} : \texttt{EOS}) \\ &\geq u(\mathbf{f} : \texttt{EOS}) - \sqrt{c \log T/N}. \end{aligned}$$

On the other hand, by DDMC, we have

$$u(\mathbf{o}^{(1:i+1)} : \texttt{EOS}) - u(\mathbf{f} : \texttt{EOS}) \leq u(\mathbf{o}^{(1:i)} : \texttt{EOS}) - u(\mathbf{a}^{(1:i)} : \texttt{EOS})$$
$$\leq 2id,$$

implying that

$$u(\mathbf{a}^{(1:i+1)} : \texttt{EOS}) \geq u(\mathbf{o}^{(1:i+1)} : \texttt{EOS}) - 2(i+1)d.$$

Thus, by the induction principle, under $\texttt{Clean}^{(1:L)}$, we have $u(\tilde{\mathbf{o}}^{(1:L)}) - u(\tilde{\mathbf{a}}^{(1:L)}) \leq 2Ld$ since we know that $\mathbf{o}^{(L)} = \mathbf{a}^{(L)} = \texttt{EOS}$ in the extended sequences.

Recall the definition of $\texttt{Clean}^{(1:L)}$:

$$\begin{aligned}\texttt{Clean}^{1:L} &= \left(\cap_{j\in[L-1]}\texttt{Clean}^{(1:L-1)}\right) \cap \left(\cap_{\tau\in\mathcal{V}}(\texttt{Clean}_{\mathbf{a}^{(1:L-1)}:\tau:\texttt{EOS}})\right) \cap \texttt{Clean}_{\mathbf{o}^{1:L-1}:\texttt{EOS}} \\ &= \prod_{t=1}^{L} \left(\cap_{\tau\in\mathcal{V}}(\texttt{Clean}_{\mathbf{a}^{(1:t-1)}:\tau:\texttt{EOS}}) \cap \texttt{Clean}_{\mathbf{o}^{1:t-1}:\texttt{EOS}}\right),\end{aligned}$$

where we abuse $\mathbf{a}^{(1:0)}$ and $\mathbf{o}^{1:0}$ to denote $\emptyset$. By the independence of the reward tapes and union bound, we have

$$\Pr\left[\cap_{\tau\in\mathcal{V}}(\texttt{Clean}_{\mathbf{a}^{(1:t-1)}:\tau:\texttt{EOS}})\right] \geq 1 - \frac{2}{T^4}\cdot n.$$

Thus, again by union bound, we have

$$\Pr\left[\,\texttt{Clean}^{(1:L)}\,\right] \geq 1 - \frac{2(n+1)L}{T^4} \geq 1 - \frac{4L}{T^4}$$

Now the entire regret can be written as:

$$\begin{aligned}
\textsc{Reg} &= \textsc{Reg}_{\text{exploration}} + \textsc{Reg}_{\text{exploitation}} \\
&\leq n \cdot L \cdot N \cdot 1 + \textsc{Reg}_{\text{exploitation}} \\
&\leq nLN + 2LTd \cdot \Pr\left[\,\texttt{Clean}^{(1:L)}\,\right] + T \cdot \Pr\left[\,\neg\texttt{Clean}^{1:L}\,\right] \\
&\leq nLN + 2LT\sqrt{\frac{2\log T}{N}} + T\frac{4nL}{T^4}.
\end{aligned}$$

Assuming that $n, L \leq T$, plugging $N = T^{2/3}(\log T)^{1/3}$ yields the regret of $\textsc{Reg} = O(nLT^{2/3}(\log T)^{1/3})$.[4]

a

Bibliography

[1] Bengt Holmstrom. On the theory of delegation. Technical report, Discussion Paper, 1980.

[2] Mark Armstrong and John Vickers. A model of delegated project choice. *Econometrica*, 78(1):213–244, 2010.

[4] If n or L is large compared with T, one can simply increase the exploration parameter in the algorithm and neutralize such dependency.

[3] Jon Kleinberg and Robert Kleinberg. Delegated search approximates efficient search. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 287–302, 2018.

[4] Amihai Glazer and Refael Hassin. Optimal contests. *Economic Inquiry*, 26(1):133–143, 1988.

[5] Ulrich Krengel and Louis Sucheston. Semiamarts and finite values. 1977.

[6] Ester Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability*, pages 1213–1216, 1984.

[7] Kiarash Banihashem, Mohammad Taghi Hajiaghayi, Piotr Krysta, and Suho Shin. Delegated choice with combinatorial constraints. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pages 275–303, 2025.

[8] Aviad Rubinstein. Beyond matroids: Secretary problem and prophet inequality with general constraints. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 324–332, 2016.

[9] Mohammad Taghi Hajiaghayi, Piotr Krysta, Mohammad Mahdavi, and Suho Shin. Delegation with costly inspection. In *Proceedings of the 26th ACM Conference on Economics and Computation*, pages 871–894, 2025.

[10] Martin Weitzman. *Optimal search for the best alternative*, volume 78. Department of Energy, 1978.

[11] Maria Arbatskaya. Ordered search. *The RAND Journal of Economics*, 38(1):119–126, 2007.

[12] MohammadTaghi Hajiaghayi, Mohammad Mahdavi, Keivan Rezaei, and Suho Shin. Regret analysis of repeated delegated choice. *arXiv preprint arXiv:2310.04884*, 2023.

[13] Suho Shin, Keivan Rezaei, and Mohammadtaghi Hajiaghayi. Delegating to multiple agents. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 1081–1126, 2023.

[14] Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2):265–281, 1983.

[15] Yuan Deng, Jieming Mao, Balasubramanian Sivan, and Kangning Wang. Approximately efficient bilateral trade. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 718–721, 2022.

[16] Jeremy I Bulow and Paul D Klemperer. Auctions vs. negotiations, 1994.

[17] Ilya Hajiaghayi, MohammadTaghi Hajiaghayi, Gary Peng, and Suho Shin. Gains-from-trade in bilateral trade with a broker. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4827–4860. SIAM, 2025.

[18] MohammadTaghi Hajiaghayi, Gary Peng, and Suho Shin. Single-sample bilateral trade with a broker. In *Proceedings of the ACM Web Conference 2026*, pages 363–374, 2026.

[19] Moshe Babaioff, Kira Goldner, and Yannai A Gonczarowski. Bulow-klemperer-style results for welfare maximization in two-sided markets. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2452–2471. SIAM, 2020.

[20] Negin Golrezaei, MohammadTaghi Hajiaghayi, and Suho Shin. The contest behind the feed: Optimal contest for recommender systems. *Available at SSRN 5258940*, 2025.

[21] A Wayne Roberts and Dale E Varberg. *Convex Functions: Convex Functions*. Academic Press, 1974.

[22] Suho Shin, Seungjoon Lee, and Jungseul Ok. Multi-armed bandit algorithm against strategic replication. In *International Conference on Artificial Intelligence and Statistics*, pages 403–431. PMLR, 2022.

[23] Suho Shin, Seyed A Esmaeili, and MohammadTaghi Hajiaghayi. Replication-proof bandit mechanism design with bayesian agents. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 20453–20460, 2025.

[24] Soheil Feizi, MohammadTaghi Hajiaghayi, Keivan Rezaei, and Suho Shin. Online advertisements with llms: Opportunities and challenges. *arXiv preprint arXiv:2311.07601*, 2023.

[25] Suho Shin, Chenghao Yang, Haifeng Xu, and MohammadTaghi Hajiaghayi. Tokenized bandit for llm decoding and alignment. In *Forty-second International Conference on Machine Learning*.

[26] MohammadTaghi Hajiaghayi, Sébastien Lahaie, Keivan Rezaei, and Suho Shin. Ad auctions for llms via retrieval augmented generation. *arXiv preprint arXiv:2406.09459*, 2024.

[27] MohammadTaghi Hajiaghayi, Sebastien Lahaie, Mohammad Mahdavi, and Suho Shin. Optimally curating an event.

[28] Robert Kleinberg, Bo Waggoner, and E Glen Weyl. Descending price optimally coordinates search. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 23–24, 2016.

[29] Hedyeh Beyhaghi and Robert Kleinberg. Pandora's problem with nonobligatory inspection. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 131–132, 2019.

[30] Yang Cai and Jinzhao Wu. On the optimal fixed-price mechanism in bilateral trade. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 737–750, 2023.

[31] Shuchi Chawla, Nikhil Devanur, and Thodoris Lykouris. Static pricing for multi-unit prophet inequalities. *arXiv preprint arXiv:2007.07990*, 2020.

[32] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47:235–256, 2002.

[33] Mohammad Taghi Hajiaghayi, Robert Kleinberg, and Tuomas Sandholm. Automated online mechanism design and prophet inequalities. In *AAAI*, volume 7, pages 58–65, 2007.

[34] Shuchi Chawla, Jason D Hartline, David L Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 311–320, 2010.

[35] Paul Dutting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing nonstochastic inputs. *SIAM Journal on Computing*, 49(3):540–582, 2020.

[36] Pablo D Azar, Robert Kleinberg, and S Matthew Weinberg. Prophet inequalities with limited information. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 1358–1377. SIAM, 2014.

[37] Hossein Esfandiari, MohammadTaghi Hajiaghayi, Vahid Liaghat, and Morteza Monemizadeh. Prophet secretary. *SIAM Journal on Discrete Mathematics*, 31(3):1685–1701, 2017.

[38] Hossein Esfandiari, MohammadTaghi HajiAghayi, Brendan Lucier, and Michael Mitzenmacher. Online pandora's boxes and bandits. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1885–1892, 2019.

[39] Saeed Alaei, MohammadTaghi Hajiaghayi, and Vahid Liaghat. Online prophet-inequality matching with applications to ad allocation. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 18–35, 2012.

[40] Michal Feldman, Nick Gravin, and Brendan Lucier. Combinatorial auctions via posted prices. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*, pages 123–135. SIAM, 2014.

[41] Melika Abolhassani, Soheil Ehsani, Hossein Esfandiari, MohammadTaghi Hajiaghayi, Robert Kleinberg, and Brendan Lucier. Beating 1-1/e for ordered prophets. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 61–71, 2017.

[42] José Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. Posted price mechanisms for a random stream of customers. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 169–186, 2017.

[43] Jose Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. Recent developments in prophet inequalities. *ACM SIGecom Exchanges*, 17(1):61–70, 2019.

[44] Robert Kleinberg and Seth Matthew Weinberg. Matroid prophet inequalities. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 123–136, 2012.

[45] Moshe Babaioff, Nicole Immorlica, and Robert Kleinberg. Matroids, secretary problems, and online mechanisms. In *Symposium on Discrete Algorithms (SODA'07)*, pages 434–443, 2007.

[46] Curtis Bechtel and Shaddin Dughmi. Delegated stochastic probing. *Innovations in Theoretical Computer Science (ITCS)*, 2021.

[47] Curtis Bechtel, Shaddin Dughmi, and Neel Patel. Delegated pandora's box. *arXiv preprint arXiv:2202.10382*, 2022.

[48] Tomer Ezra, Michal Feldman, Nick Gravin, and Zhihao Gavin Tang. "who is next in line?" on the significance of knowing the arrival order in bayesian online settings. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3759–3776. SIAM, 2023.

[49] Tomer Ezra, Michal Feldman, and Zhihao Gavin Tang. Choosing behind the veil: Tight bounds for identity-blind online algorithms. *arXiv preprint arXiv:2402.17160*, 2024.

[50] Xinye Li and Andrew Chi-Chih Yao. On revenue maximization for selling multiple independently distributed items. *Proceedings of the National Academy of Sciences*, 110(28):11232–11237, 2013.

[51] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. A sharp concentration inequality with applications. *Random Structures & Algorithms*, 16(3):277–292, 2000.

[52] Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 1014–1033. SIAM, 2016.

[53] Shuchi Chawla, Kira Goldner, Anna R Karlin, and J Benjamin Miller. Non-adaptive matroid prophet inequalities. *arXiv preprint arXiv:2011.09406*, 2020.

[54] Alice Sayutina. Non-adaptive matroid prophet inequalities. Master's thesis, University of Waterloo, 2023.

[55] Laura Doval. Whether or not to open pandora's box. *Journal of Economic Theory*, 175:127–158, 2018.

[56] Hu Fu, Jiawei Li, and Daogao Liu. Pandora box problem with nonobligatory inspection: Hardness and approximation scheme. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 789–802, 2023.

[57] Ben Berger, Tomer Ezra, Michal Feldman, and Federico Fusco. Pandora's problem with combinatorial cost. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 273–292, 2023.

[58] Hedyeh Beyhaghi and Linda Cai. Pandora's problem with nonobligatory inspection: Optimal structure and a ptas. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 803–816, 2023.

[59] Ricardo Alonso and Niko Matouschek. Optimal delegation. *The Review of Economic Studies*, 75(1):259–293, 2008.

[60] Ali Khodabakhsh, Emmanouil Pountourakis, and Samuel Taggart. Simple delegated choice. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 569–590. SIAM, 2024.

[61] Curtis Bechtel and Shaddin Dughmi. Delegated stochastic probing. *arXiv preprint arXiv:2010.14718*, 2020.

[62] Mohammad Hajiaghayi, Mohammad Mahdavi, Keivan Rezaei, and Suho Shin. Regret analysis of repeated delegated choice. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 9757–9764, 2024.

[63] Sudipto Guha, Kamesh Munagala, and Saswati Sarkar. Information acquisition and exploitation in multichannel wireless networks. *arXiv preprint arXiv:0804.1724*, 2008.

[64] H Fu, J Li, and P Xu. A ptas for a class of stochastic dynamic programs. chatzigiannakis i, kaklamanis c, marx d, sannella d, eds. In *Proc. 45th Internat. Colloquium on Automata, Languages, and Programming.(ICALP 2018), Prague, Czech Republic*, pages 1–56, 2018.

[65] Sahil Singla. The price of information in combinatorial optimization. In *Proceedings of the twenty-ninth annual ACM-SIAM symposium on discrete algorithms*, pages 2523–2532. SIAM, 2018.

[66] Robin Bowers and Bo Waggoner. Matching with nested and bundled pandora boxes. *arXiv preprint arXiv:2406.08711*, 2024.

[67] Ulrich Krengel and Louis Sucheston. Prophet compared to gambler: an inequality for transforms of processes. *The Annals of Probability*, 15(4):1593–1599, 1987.

[68] Tan Gan, Ju Hu, and Xi Weng. Optimal contingent delegation. *Journal of Economic Theory*, page 105597, 2022.

[69] Ricardo Alonso, Isabelle Brocas, and Juan D Carrillo. Resource allocation in the brain. *Review of Economic Studies*, 81(2):501–534, 2014.

[70] Thomas W Gilligan and Keith Krehbiel. Asymmetric information and legislative rules with a heterogeneous committee. *American journal of political science*, pages 459–490, 1989.

[71] Vijay Krishna. Asymmetric information and legislative rules: Some amendments. *American Political science review*, 95(2):435–452, 2001.

[72] David Martimort and Aggey Semenov. The informational effects of competition and collusion in legislative politics. *Journal of Public Economics*, 92(7):1541–1563, 2008.

[73] William Fuchs, Satoshi Fukuda, and Mahyar Sefidgaran. Who to listen to?: A model of endogenous delegation. 2022.

[74] Robert Ulbricht. Optimal delegated search with adverse selection and moral hazard. *Theoretical Economics*, 11(1):253–278, 2016.

[75] Vijay Krishna and John Morgan. Contracting for information under imperfect commitment. *The RAND Journal of Economics*, 39(4):905–925, 2008.

[76] Tracy R Lewis. A theory of delegated search for the best alternative. *The RAND Journal of Economics*, 43(3):391–416, 2012.

[77] Xiaohu Li. A note on expected rent in auction theory. *Operations Research Letters*, 33(5):531–534, 2005.

[78] Mitchell Watt. Concavity and convexity of order statistics in sample size. *arXiv preprint arXiv:2111.04702*, 2021.

[79] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.

[80] Edward P Lazear and Sherwin Rosen. Rank-order tournaments as optimum labor contracts. *Journal of political Economy*, 89(5):841–864, 1981.

[81] Yasar Barut and Dan Kovenock. The symmetric multiple prize all-pay auction with complete information. *European Journal of Political Economy*, 14(4):627–644, 1998.

[82] Benny Moldovanu and Aner Sela. The optimal allocation of prizes in contests. In *40 Years of Research on Rent Seeking 1*, pages 615–631. Springer, 2008.

[83] Dawei Fang, Thomas Noe, and Philipp Strack. Turning up the heat: The discouraging effect of competition in contests. *Journal of Political Economy*, 128(5):1940–1975, 2020.

[84] Arpita Ghosh and Patrick Hummel. A game-theoretic analysis of rank-order mechanisms for user-generated content. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 189–198, 2011.

[85] Arpita Ghosh and Robert Kleinberg. Optimal contest design for simple agents. *ACM Transactions on Economics and Computation (TEAC)*, 4(4):1–41, 2016.

[86] Amy Greenwald, Takehiro Oyakawa, and Vasilis Syrgkanis. Simple vs optimal contests with convex costs. In *Proceedings of the 2018 World Wide Web Conference*, pages 1429–1438, 2018.

[87] Shuchi Chawla, Jason D Hartline, and Balasubramanian Sivan. Optimal crowdsourcing contests. *Games and Economic Behavior*, 113:80–96, 2019.

[88] Philip J Reny. On the existence of pure and mixed strategy nash equilibria in discontinuous games. *Econometrica*, 67(5):1029–1056, 1999.

[89] Meena Jagadeesan, Nikhil Garg, and Jacob Steinhardt. Supply-side equilibria in recommender systems. *Advances in Neural Information Processing Systems*, 36, 2024.

[90] Samuel Karlin. Total positivity, absorption probabilities and applications. *Transactions of the American Mathematical Society*, 111(1):33–107, 1964.

[91] Shang-jun Yang, Hua-zhang Wu, and Quan-bing Zhang. Generalization of vandermonde determinants. *Linear Algebra and its Applications*, 336(1-3):201–204, 2001.

[92] Dmitrii Karp, Anna Vishnyakova, and Yi Zhang. Unimodality preservation by ratios of functional series and integral transforms. *arXiv preprint arXiv:2408.01755*, 2024.

[93] François-Désiré Budan. *Nouvelle méthode pour la résolution des équations numériques d'un degré quelconque: d'après laquelle toute le calcul exigé pour cette résolution se réduit à l'emploi des deux premières règles de l'arithmétique*. Courcier, 1807.

[94] Xiaoshen Wang. A simple proof of descartes's rule of signs. *The American Mathematical Monthly*, 111(6):525–526, 2004.

[95] Nicholas C Barberis. Thirty years of prospect theory in economics: A review and assessment. *Journal of economic perspectives*, 27(1):173–196, 2013.

[96] Dana Sisak. Multiple-prize contests–the optimal allocation of prizes. *Journal of Economic Surveys*, 23(1):82–114, 2009.

[97] Nicole Immorlica, Meena Jagadeesan, and Brendan Lucier. Clickbait vs. quality: How engagement-based optimization shapes the content landscape in online platforms. *arXiv preprint arXiv:2401.09804*, 2024.

[98] Meena Jagadeesan, Michael I Jordan, and Nika Haghtalab. Competition, alignment, and equilibria in digital marketplaces. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 5689–5696, 2023.

[99] Arpita Ghosh and Preston McAfee. Incentivizing high-quality user-generated content. In *Proceedings of the 20th international conference on World wide web*, pages 137–146, 2011.

[100] Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. How bad is top-k recommendation under competing content creators? In *International Conference on Machine Learning*, pages 39674–39701. PMLR, 2023.

[101] Fan Yao, Chuanhao Li, Karthik Abinav Sankararaman, Yiming Liao, Yan Zhu, Qifan Wang, Hongning Wang, and Haifeng Xu. Rethinking incentives in recommender systems: are monotone rewards always beneficial? *Advances in Neural Information Processing Systems*, 36, 2024.

[102] Josip E Peajcariaac and Yung Liang Tong. *Convex functions, partial orderings, and statistical applications*. Academic Press, 1992.

[103] Aleksandrs Slivkins et al. Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2):1–286, 2019.

[104] Andrea Barraza-Urbina and Dorota Glowacka. Introduction to bandits in recommender systems. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 748–750, 2020.

[105] Liang Tang, Yexi Jiang, Lei Li, and Tao Li. Ensemble contextual bandits for personalized recommendation. In *Proceedings of the 8th ACM Conference on Recommender Systems*, pages 73–80, 2014.

[106] Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. Bandits with knapsacks. *Journal of the ACM (JACM)*, 65(3):1–55, 2018.

[107] Ashwinkumar Badanidiyuru, John Langford, and Aleksandrs Slivkins. Resourceful contextual bandits. In *Conference on Learning Theory*, pages 1109–1134. PMLR, 2014.

[108] Bouneffouf Djallel and Rish Irina. A survey on practical applications of multi-armed and contextual bandits. *arXiv preprint arXiv:1904.10040*, 2019.

[109] Lydia T Liu, Feng Ruan, Horia Mania, and Michael I Jordan. Bandit learning in decentralized matching markets. *The Journal of Machine Learning Research*, 22(1):9612–9645, 2021.

[110] Lydia T Liu, Horia Mania, and Michael Jordan. Competing bandits in matching markets. In *International Conference on Artificial Intelligence and Statistics*, pages 1618–1628. PMLR, 2020.

[111] Mark Braverman, Jieming Mao, Jon Schneider, and S Matthew Weinberg. Multi-armed bandit problems with strategic arms. In *Conference on Learning Theory*, pages 383–416. PMLR, 2019.

[112] Zhe Feng, David Parkes, and Haifeng Xu. The intrinsic robustness of stochastic bandits to strategic manipulation. In *International Conference on Machine Learning*, pages 3092–3101. PMLR, 2020.

[113] Seyed Esmaeili, MohammadTaghi Hajiaghayi, and Suho Shin. Replication-proof bandit mechanism design. *arXiv preprint arXiv:2312.16896*, 2023.

[114] Jason D Hartline et al. Bayesian mechanism design. *Foundations and Trends® in Theoretical Computer Science*, 8(3):143–263, 2013.

[115] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge university press, 2007.

[116] Shipra Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *Conference on learning theory*, pages 39–1. JMLR Workshop and Conference Proceedings, 2012.

[117] Jean-Yves Audibert, Sébastien Bubeck, and Rémi Munos. Best arm identification in multi-armed bandits. In *COLT*, pages 41–53, 2010.

[118] Aurélien Garivier, Tor Lattimore, and Emilie Kaufmann. On explore-then-commit strategies. *Advances in Neural Information Processing Systems*, 29, 2016.

[119] Yaming Yu. Stochastic ordering of exponential family distributions and their mixturesxk. *Journal of Applied Probability*, 46(1):244–254, 2009.

[120] R Preston McAfee. The gains from trade under fixed price mechanisms. *Applied economics research bulletin*, 1(1):1–10, 2008.

[121] Liad Blumrosen and Shahar Dobzinski. Reallocation mechanisms. *arXiv preprint arXiv:1404.6786*, 2014.

[122] Liad Blumrosen and Shahar Dobzinski. (almost) efficient mechanisms for bilateral trading. *arXiv preprint arXiv:1604.04876*, 2016.

[123] Liad Blumrosen and Yehonatan Mizrahi. Approximating gains-from-trade in bilateral trading. In *Web and Internet Economics: 12th International Conference, WINE 2016, Montreal, Canada, December 11-14, 2016, Proceedings 12*, pages 400–413. Springer, 2016.

[124] Jernej Čopič and Clara Ponsatí. Robust bilateral trade and mediated bargaining. *Journal of the European Economic Association*, 6(2-3):570–580, 2008.

[125] Xingtan Zhang. *Efficient bargaining through a broker*. 2019.

[126] Ran Eilat and Ady Pauzner. Bilateral trade with a benevolent intermediary. *Theoretical Economics*, 16(4):1655–1714, 2021.

[127] Zhonghong Kuang, Weiran Shen, and Fan Wu. Profit-maximizing mechanism in bilateral trade with interdependent valuations. *Available at SSRN 4474002*, 2023.

[128] Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

[129] Zi Yang Kang and Jan Vondrák. Fixed-price approximations to optimal efficiency in bilateral trade. *Available at SSRN 3460336*, 2019.

[130] Hans Degryse, Mark Van Achter, and Gunther Wuyts. Dynamic order submission strategies with competition between a dealer market and a crossing network. *Journal of Financial Economics*, 91(3):319–338, 2009.

[131] Haoxiang Zhu. Do dark pools harm price discovery? *The Review of Financial Studies*, 27(3):747–789, 2014.

[132] Darrell Duffie and Haoxiang Zhu. Size discovery. *The Review of Financial Studies*, 30(4):1095–1150, 2017.

[133] Kathleen M Hagerty and William P Rogerson. Robust trading mechanisms. *Journal of Economic Theory*, 42(1):94–107, 1987.

[134] Jason D Hartline and Tim Roughgarden. Simple versus optimal mechanisms. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 225–234, 2009.

[135] Shuchi Chawla, Jason D Hartline, and Robert Kleinberg. Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, pages 243–251, 2007.

[136] Jason D Hartline and Tim Roughgarden. Optimal mechanism design and money burning. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 75–84, 2008.

[137] Tim Roughgarden. Algorithmic game theory. *Communications of the ACM*, 53(7):78–86, 2010.

[138] Yumou Fei. Improved approximation to first-best gains-from-trade. In *International Conference on Web and Internet Economics*, pages 204–218. Springer, 2022.

[139] Shahar Dobzinski and Ariel Shaulker. Bilateral trade with correlated values. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 237–246, 2024.

[140] Zhengyang Liu, Zeyu Ren, and Zihe Wang. Improved approximation ratios of fixed-price mechanisms in bilateral trades. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 751–760, 2023.

[141] Zi Yang Kang, Francisco Pernice, and Jan Vondrák. Fixed-price approximations in bilateral trade. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2964–2985. SIAM, 2022.

[142] R Preston McAfee. A dominant strategy double auction. *Journal of economic Theory*, 56(2):434–450, 1992.

[143] Kaustubh Deshmukh, Andrew V Goldberg, Jason D Hartline, and Anna R Karlin. Truthful and competitive double auctions. In *Algorithms—ESA 2002: 10th Annual European Symposium Rome, Italy, September 17–21, 2002 Proceedings 10*, pages 361–373. Springer, 2002.

[144] Sandeep Baliga and Rakesh Vohra. Market research and market design. *Advances in theoretical Economics*, 3(1), 2003.

[145] Yang Cai, Christopher Liaw, Aranyak Mehta, and Mingfei Zhao. The power of two-sided recruitment in two-sided markets. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 201–212, 2024.

[146] Johannes Brustle, Yang Cai, Fa Wu, and Mingfei Zhao. Approximating gains from trade in two-sided markets via simple mechanisms. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 589–590, 2017.

[147] Yang Cai, Kira Goldner, Steven Ma, and Mingfei Zhao. On multi-dimensional gains from trade maximization. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1079–1098. SIAM, 2021.

[148] Ilya Hajiaghayi, MohammadTaghi Hajiaghayi, Gary Peng, and Suho Shin. Gains-from-trade in bilateral trade with a broker. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms*, 2025.

[149] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. In *Proceedings of the 11th ACM conference on Electronic commerce*, pages 129–138, 2010.

[150] Eric Maskin and John Riley. Asymmetric auctions. *The review of economic studies*, 67(3):413–438, 2000.

[151] René Kirkegaard. A mechanism design approach to ranking asymmetric auctions. *Econometrica*, 80(5):2349–2364, 2012.

[152] Zhiyi Huang, Yishay Mansour, and Tim Roughgarden. Making the most of your samples. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 45–60, 2015.

[153] Hu Fu, Nicole Immorlica, Brendan Lucier, and Philipp Strack. Randomization beats second price as a prior-independent auction. In *Proceedings of the 16th ACM Conference on Economics and Computation*, 2015.

[154] Kira Goldner and Anna R. Karlin. A prior-independent revenue-maximizing auction for multiple additive bidders. In *Proceedings of the 12th International Conference on Web and Internet Economics*, 2016.

[155] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[156] Romal Thoppilan, Daniel De Freitas, Jamie Hall, Noam Shazeer, Apoorv Kulshreshtha, Heng-Tze Cheng, Alicia Jin, Taylor Bos, Leslie Baker, Yu Du, et al. Lamda: Language models for dialog applications. *arXiv preprint arXiv:2201.08239*, 2022.

[157] Daniel Fried, Armen Aghajanyan, Jessy Lin, Sida Wang, Eric Wallace, Freda Shi, Ruiqi Zhong, Wen-tau Yih, Luke Zettlemoyer, and Mike Lewis. Incoder: A generative model for code infilling and synthesis. *arXiv preprint arXiv:2204.05999*, 2022.

[158] Yizhen Zheng, Huan Yee Koh, Maddie Yang, Li Li, Lauren T May, Geoffrey I Webb, Shirui Pan, and George Church. Large language models in drug discovery and development: From disease mechanisms to clinical trials. *arXiv preprint arXiv:2409.04481*, 2024.

[159] Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, et al. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.

[160] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.

[161] Swaroop Mishra, Daniel Khashabi, Chitta Baral, and Hannaneh Hajishirzi. Cross-task generalization via natural language crowdsourcing instructions. *arXiv preprint arXiv:2104.08773*, 2021.

[162] Victor Sanh, Albert Webson, Colin Raffel, Stephen H Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Teven Le Scao, Arun Raja, et al. Multitask prompted training enables zero-shot task generalization. *arXiv preprint arXiv:2110.08207*, 2021.

[163] Tianhao Shen, Renren Jin, Yufei Huang, Chuang Liu, Weilong Dong, Zishan Guo, Xinwei Wu, Yan Liu, and Deyi Xiong. Large language model alignment: A survey. *arXiv preprint arXiv:2309.15025*, 2023.

[164] Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*, 2022.

[165] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35:27730–27744, 2022.

[166] Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in Neural Information Processing Systems*, 36, 2024.

[167] Zhichao Wang, Bin Bi, Shiva Kumar Pentyala, Kiran Ramnath, Sougata Chaudhuri, Shubham Mehrotra, Xiang-Bo Mao, Sitaram Asur, et al. A comprehensive survey of llm alignment techniques: Rlhf, rlaif, ppo, dpo and more. *arXiv preprint arXiv:2407.16216*, 2024.

[168] Martin Josifoski, Maxime Peyrard, Frano Rajic, Jiheng Wei, Debjit Paul, Valentin Hartmann, Barun Patra, Vishrav Chaudhary, Emre Kıcıman, Boi Faltings, et al. Language model decoding as likelihood-utility alignment. *arXiv preprint arXiv:2210.07228*, 2022.

[169] James Y Huang, Sailik Sengupta, Daniele Bonadiman, Yi-an Lai, Arshit Gupta, Nikolaos Pappas, Saab Mansour, Katrin Kirchhoff, and Dan Roth. Deal: Decoding-time alignment for large language models. *arXiv preprint arXiv:2402.06147*, 2024.

[170] Sidharth Mudgal, Jong Lee, Harish Ganapathy, YaGuang Li, Tao Wang, Yanping Huang, Zhifeng Chen, Heng-Tze Cheng, Michael Collins, Trevor Strohman, et al. Controlled decoding from language models. *arXiv preprint arXiv:2310.17022*, 2023.

[171] Souradip Chakraborty, Soumya Suvra Ghosal, Ming Yin, Dinesh Manocha, Mengdi Wang, Amrit Singh Bedi, and Furong Huang. Transfer q star: Principled decoding for llm alignment. *arXiv preprint arXiv:2405.20495*, 2024.

[172] Bolian Li, Yifan Wang, Ananth Grama, and Ruqi Zhang. Cascade reward sampling for efficient decoding-time alignment. *arXiv preprint arXiv:2406.16306*, 2024.

[173] Ruizhe Shi, Yifang Chen, Yushi Hu, Alisa Liu, Hannaneh Hajishirzi, Noah A Smith, and Simon S Du. Decoding-time language model alignment with multiple objectives. *arXiv preprint arXiv:2406.18853*, 2024.

[174] Matthew Finlayson, John Hewitt, Alexander Koller, Swabha Swayamdipta, and Ashish Sabharwal. Closing the curious case of neural text degeneration. *arXiv preprint arXiv:2310.01693*, 2023.

[175] Sijin Chen, Omar Hagrass, and Jason M Klusowski. Decoding game: On minimax optimality of heuristic text generation strategies. *arXiv preprint arXiv:2410.03968*, 2024.

[176] Pierre-Arnaud Coquelin and Rémi Munos. Bandit algorithms for tree search. *arXiv preprint cs/0703062*, 2007.

[177] Yifan Song, Guoyin Wang, Sujian Li, and Bill Yuchen Lin. The good, the bad, and the greedy: Evaluation of llms should not ignore non-determinism. *arXiv preprint arXiv:2407.10457*, 2024.

[178] Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. *Advances in neural information processing systems*, 24, 2011.

[179] Ximing Lu, Sean Welleck, Peter West, Liwei Jiang, Jungo Kasai, Daniel Khashabi, Ronan Le Bras, Lianhui Qin, Youngjae Yu, Rowan Zellers, et al. Neurologic a* esque decoding: Constrained text generation with lookahead heuristics. *arXiv preprint arXiv:2112.08726*, 2021.

[180] Yichao Fu, Peter Bailis, Ion Stoica, and Hao Zhang. Break the sequential dependency of llm inference using lookahead decoding. *arXiv preprint arXiv:2402.02057*, 2024.

[181] Clara Meister, Tim Vieira, and Ryan Cotterell. If beam search is the answer, what was the question? *arXiv preprint arXiv:2010.02650*, 2020.

[182] Chufan Shi, Haoran Yang, Deng Cai, Zhisong Zhang, Yifan Wang, Yujiu Yang, and Wai Lam. A thorough examination of decoding methods in the era of llms. *arXiv preprint arXiv:2402.06925*, 2024.

[183] Sean Welleck, Amanda Bertsch, Matthew Finlayson, Hailey Schoelkopf, Alex Xie, Graham Neubig, Ilia Kulikov, and Zaid Harchaoui. From decoding to meta-generation: Inference-time algorithms for large language models. *arXiv preprint arXiv:2406.16838*, 2024.

[184] Pan Li and Olgica Milenkovic. Inhomogeneous hypergraph clustering with applications. *Advances in neural information processing systems*, 30, 2017.

[185] Sebastian Tschiatschek, Adish Singla, and Andreas Krause. Selecting sequences of items via submodular maximization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

[186] Marko Mitrovic, Moran Feldman, Andreas Krause, and Amin Karbasi. Submodularity on hypergraphs: From sets to sequences. In *International Conference on Artificial Intelligence and Statistics*, pages 1177–1184. PMLR, 2018.

[187] Marko Mitrovic, Ehsan Kazemi, Moran Feldman, Andreas Krause, and Amin Karbasi. Adaptive sequence submodularity. *Advances in Neural Information Processing Systems*, 32, 2019.

[188] Markus Freitag and Yaser Al-Onaizan. Beam search strategies for neural machine translation. *arXiv preprint arXiv:1702.01806*, 2017.

[189] Ilia Kulikov, Alexander H Miller, Kyunghyun Cho, and Jason Weston. Importance of search and evaluation strategies in neural dialogue modeling. *arXiv preprint arXiv:1811.00907*, 2018.

[190] Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*, 2019.

[191] Shicong Cen, Jincheng Mei, Katayoon Goshvadi, Hanjun Dai, Tong Yang, Sherry Yang, Dale Schuurmans, Yuejie Chi, and Bo Dai. Value-incentivized preference optimization: A unified approach to online and offline rlhf. *arXiv preprint arXiv:2405.19320*, 2024.

[192] Joy Qiping Yang, Salman Salamatian, Ziteng Sun, Ananda Theertha Suresh, and Ahmad Beirami. Asymptotics of language model alignment. *arXiv preprint arXiv:2404.01730*, 2024.

[193] Kiho Park, Yo Joong Choe, and Victor Veitch. The linear representation hypothesis and the geometry of large language models. *arXiv preprint arXiv:2311.03658*, 2023.

[194] Bernhard H Korte, Jens Vygen, B Korte, and J Vygen. *Combinatorial optimization*, volume 1. Springer, 2011.

[195] Lihong Li, Wei Chu, John Langford, and Robert E Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 661–670, 2010.

[196] Alexandre Rame, Guillaume Couairon, Corentin Dancette, Jean-Baptiste Gaya, Mustafa Shukor, Laure Soulier, and Matthieu Cord. Rewarded soups: towards pareto-optimal alignment by interpolating weights fine-tuned on diverse rewards. *Advances in Neural Information Processing Systems*, 36, 2024.

[197] Sadat Shahriar, Zheng Qi, Nikolaos Pappas, Srikanth Doss, Monica Sunkara, Kishaloy Halder, Manuel Mager, and Yassine Benajiba. Inference time llm alignment in single and multidomain preference spectrum. *arXiv preprint arXiv:2410.19206*, 2024.

[198] Joel Jang, Seungone Kim, Bill Yuchen Lin, Yizhong Wang, Jack Hessel, Luke Zettlemoyer, Hannaneh Hajishirzi, Yejin Choi, and Prithviraj Ammanabrolu. Personalized soups: Personalized large language model alignment via post-hoc parameter merging. *arXiv preprint arXiv:2310.11564*, 2023.

[199] Stephanie Lin, Jacob Hilton, and Owain Evans. Truthfulqa: Measuring how models mimic human falsehoods. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3214–3252, 2022.

[200] Saeed Alaei, Ali Makhdoumi, and Azarakhsh Malekian. Maximizing sequence-submodular functions and its application to online advertising. *Management Science*, 67(10):6030–6054, 2021.

[201] Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. In *International Conference on Learning Representations*, 2020.

[202] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*, 2022.

[203] AI@Meta. Llama 3 model card. 2024.

[204] James Oxley. Matroid theory. In *Handbook of the Tutte Polynomial and Related Topics*, pages 44–85. Chapman and Hall/CRC, 2022.

[205] J v. Neumann. Zur theorie der gesellschaftsspiele. *Mathematische annalen*, 100(1):295–320, 1928.

[206] Dirk Bergemann and Stephen Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522, 2016.

[207] Manuel Lopez and James Marengo. An upper bound for the expected difference between order statistics. *Mathematics Magazine*, 84(5):365–369, 2011.

[208] HA David. Augmented order statistics and the biasing effect of outliers. *Statistics & probability letters*, 36(2):199–204, 1997.

[209] Richard E Barlow and Frank Proschan. Statistical theory of reliability and life testing: probability models. Technical report, Florida State Univ Tallahassee, 1975.

[210] Pietro Cerone and Sever S Dragomir. Bounds for the gini mean difference via the sonin identity. *Computers & Mathematics with Applications*, 50(3-4):599–609, 2005.

[211] Sam Elder. Bayesian adaptive data analysis guarantees from subgaussianity. *arXiv preprint arXiv:1611.00065*, 2016.

[212] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2016.

[213] Seyed A Esmaeili, Suho Shin, and Aleksandrs Slivkins. Robust and performance incentivizing algorithms for multi-armed bandits with strategic agents. *arXiv preprint arXiv:2312.07929*, 2023.

[214] Arpita Ghosh and Patrick Hummel. Learning and incentives in user-generated content: Multi-armed bandits with endogenous arms. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 233–246, 2013.

[215] Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the "wisdom of the crowd". *Journal of Political Economy*, 122(5):988–1012, 2014.

[216] Yeon-Koo Che and Johannes Hörner. Recommender systems as mechanisms for social learning. *The Quarterly Journal of Economics*, 133(2):871–925, 2018.

[217] Yishay Mansour, Aleksandrs Slivkins, and Vasilis Syrgkanis. Bayesian incentive-compatible bandit exploration. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 565–582, 2015.

[218] Nicole Immorlica, Jieming Mao, Aleksandrs Slivkins, and Zhiwei Steven Wu. Incentivizing exploration with selective data disclosure. *arXiv preprint arXiv:1811.06026*, 2018.

[219] Mark Sellke and Aleksandrs Slivkins. The price of incentivizing exploration: A characterization via thompson sampling and sample complexity. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 795–796, 2021.

[220] Peter Frazier, David Kempe, Jon Kleinberg, and Robert Kleinberg. Incentivizing exploration. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 5–22, 2014.

[221] Bangrui Chen, Peter Frazier, and David Kempe. Incentivizing exploration by heterogeneous users. In *Conference On Learning Theory*, pages 798–818. PMLR, 2018.

[222] Abhijit V Banerjee. A simple model of herd behavior. *The quarterly journal of economics*, 107(3):797–817, 1992.

[223] Venkatesh Bala and Sanjeev Goyal. Learning from neighbours. *The review of economic studies*, 65(3):595–621, 1998.

[224] David Lazer and Allan Friedman. The network structure of exploration and exploitation. *Administrative science quarterly*, 52(4):667–694, 2007.

[225] Kiarash Banihashem, MohammadTaghi Hajiaghayi, Suho Shin, and Aleksandrs Slivkins. Bandit social learning: Exploration under myopic behavior. *arXiv preprint arXiv:2302.07425*, 2023.

[226] Google AdSense. Youtube adsense replication policies. `https://www.youtube.com/watch?v=F-sX_XRdtTM`, 2020.

[227] Google. Ad replication. `https://www.youtube.com/watch?v=F-sX_XRdtTM`, 2022.

[228] Fabio A Drucker and Lisa K Fleischer. Simpler sybil-proof mechanisms for multi-level marketing. In *Proceedings of the 13th ACM conference on Electronic commerce*, pages 441–458, 2012.

[229] Wei Chen, Yajun Wang, Dongxiao Yu, and Li Zhang. Sybil-proof mechanisms in query incentive networks. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 197–214, 2013.

[230] John L Kelley. *General topology*. Courier Dover Publications, 2017.

[231] Kalyanapuram Rangachari Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., 2005.

[232] Lawrence Narici and Edward Beckenstein. *Topological vector spaces*. Chapman and Hall/CRC, 2010.

[233] Andreu Mas-Colell, Michael Dennis Whinston, Jerry R Green, et al. *Microeconomic theory*, volume 1. Oxford university press New York, 1995.

[234] George Casella and Roger Berger. *Statistical inference*. CRC press, 2024.

[235] George G Lorentz. *Bernstein polynomials*. American Mathematical Soc., 2012.

[236] EH Doha, AH Bhrawy, and MA Saker. On the derivatives of bernstein polynomials: an application for the solution of high even-order differential equations. *Boundary Value Problems*, 2011:1–16, 2011.

[237] Wenchang Chu. Abel's lemma on summation by parts and basic hypergeometric series. *Advances in Applied Mathematics*, 39(4):490–514, 2007.

[238] Isaac J Schoenberg. On variation diminishing approximation methods. *On numerical approximation*, 2:249–274, 1959.

[239] David Raymond Curtiss. Recent extentions of descartes' rule of signs. *Annals of Mathematics*, 19(4):251–278, 1918.

[240] Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

[241] Ahmad Beirami, Alekh Agarwal, Jonathan Berant, Alexander D'Amour, Jacob Eisenstein, Chirag Nagpal, and Ananda Theertha Suresh. Theoretical guarantees on the best-of-n alignment policy. *arXiv preprint arXiv:2401.01879*, 2024.

[242] Sourya Basu, Govardana Sachitanandam Ramachandran, Nitish Shirish Keskar, and Lav R Varshney. Mirostat: A neural text decoding algorithm that directly controls perplexity. *arXiv preprint arXiv:2007.14966*, 2020.

[243] Andrew Chi-Chin Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, pages 222–227. IEEE Computer Society, 1977.