INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2025IPPAG005

Thèse de doctorat

The Impact of Fairness in Resource Allocation

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à l'École nationale de la statistique et de l'administration économique

École doctorale n°574 École doctorale de mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Palaiseau, le 23 Septembre 2025, par

MATHIEU MOLINA

Composition du Jury :

Jean-Michel Loubès Directeur de recherche, Inria & ANITI	Rapporteur
Omar Besbes Professeur, Columbia University	Rapporteur
Laurent Massoulié Directeur de recherche, Inria	Examinateur
Olivier Cappé Directeur de recherche, CNRS	Examinateur
Azadeh Khaleghi Professeure, ENSAE	Examinatrice
Vianney Perchet Professeur, ENSAE	Directeur de thèse
Patrick Loiseau Directeur de recherche, Inria	Co-directeur de thèse
Nicolas Gast Chargé de recherche, Inria	Co-directeur de thèse

Abstract

Resource allocation problems, ranging from school choice to online advertising and healthcare, traditionally prioritize efficiency and social welfare while often overlooking ethical concerns. Introducing fairness constraints into well-established allocation problems requires new algorithmic design principles, alters performance, and gives rise to new trade-offs between optimality and constraints satisfaction. This thesis studies four standard allocation models – online learning, prophet inequalities, auctions, and matroid-based allocation – with equity considerations, and explores how enforcing fairness impacts the allocation quality through regret bounds, competitive ratio, inequality measures, and price of fairness. We show that, in some cases, trade-offs are unavoidable; in others, the loss becomes arbitrarily small asymptotically; and in specific instances, fairness can be achieved at no cost. By characterizing when and how fairness constraints affect allocation outcomes, this thesis provides guidance to both algorithm designers and policy makers seeking to deploy fair algorithms in some allocation problem scenarios.

Remerciements

Je voudrais tout d'abord remercier mes directeurs de thèse. Patrick, je me rappelle de t'avoir contacté, il y a plus de trois ans, encore bien incertain sur mon avenir. Aujourd'hui, grâce à Vianney et Nicolas qui ont accepté, à mon grand bonheur de m'encadrer, j'ai trouvé une voie qui me passionne, et que j'espère poursuivre encore longtemps. Et pour cela, merci de m'avoir fait confiance en me prenant en thèse. J'aurais pu vous dire que j'ai vécu des moments difficiles durant cette thèse, que j'ai finalement pu surmonter héroïquement, mais ce n'est pas le cas. Vianney, Patrick, Nicolas, votre bonne humeur et vos saillies durant nos meetings hebdomadaires, parfois un peu taquines, ont rendu ces réunions toujours agréables. Travailler avec vous a été un vrai plaisir, et j'ai partagé avec vous de nombreux moments précieux. Je reste toujours impressionné par la profondeur de vos connaissances, vos capacités, vos qualités humaines, et j'ai beaucoup de respect pour vous. Même si je peux parfois être un peu têtu, vos conseils ont toujours compté pour moi et m'ont fait grandir. J'ai tant appris de vous, je me rends compte de la chance que j'ai eue et je vous en suis profondément reconnaissant. Merci du fond du cœur.

Je tiens à remercier chaleureusement Omar Besbes et Jean-Michel Loubès d'avoir accepté de relire mon manuscrit et d'y avoir consacré du temps, ainsi que pour leurs rapports détaillés, encourageants et constructifs. Je suis également très reconnaissant à Olivier Cappé, Azadeh Khaleghi et Laurent Massoulié d'avoir accepté de faire partie de mon jury : c'est pour moi un grand honneur de pouvoir partager mes travaux avec vous.

Un grand merci à Marco et Xavier, pour votre accueil si chaleureux à Rome. Ce fut un plaisir d'échanger avec vous, et votre soutien m'a été d'une grande valeur. Le seul reproche que je pourrais vous faire, c'est d'avoir rendu mes attentes encore plus exigeantes pour la cuisine italienne! Merci aussi aux membres de LUISS, Maurizio, Michalis, Marta, Matteo, Roberto, Emilio, ainsi qu'à tous les chercheurs que j'ai eu la chance de rencontrer ou de retrouver en Italie, qui m'ont fait aimer ce pays : Johannes, Andrea, Stefano, Matteo, Federico, Tomasz et bien d'autres.

Je voudrais exprimer ma gratitude à Mauricio, grâce à qui j'ai eu mon premier contact avec la recherche. Ton stage a été pour moi une expérience déterminante, et tes conseils m'ont convaincu de tenter cette aventure. Merci aussi à Émilie, qui m'a toujours encouragée à faire une thèse et dont les enseignements m'ont particulièrement formateurs.

Je veux remercier de tout cœur tous mes co-auteurs, avec qui j'ai passé tant d'heures au tableau, en visioconférence ou simplement autour d'un (des?) verre. Merci à Bary, Hugo, Dorian, Simon M, Simon F, Flore, Felipe, Rémi, Nadav, Victor, Sebastian et Corentin. Je fonctionne à l'échange d'idées, et sans vous, rien n'aurait été possible. J'ai toujours eu grand plaisir à discuter avec vous : ce furent autant de moments intellectuellement enrichissants que de moments humains qui me sont chers, et j'espère que ces collaborations pourront se prolonger encore longtemps.

L'équipe (au sens large) de FAIRPLAY, où j'ai passé l'entièreté de ma thèse, a été simplement formidable. Si je venais chaque semaine le sourire aux lèvres, c'était grâce à vous et à votre bonne humeur collective. Merci en particulier à Lorenzo, Marius, Solenne, Ziyad, Côme, Mike, Hafedh, Réda, Rémi, Matilde, Giovanni, Corentin, Marie, Minrui, Shiyun, Mélissa, Étienne, Evrard, Sasila, Denis, Bartholomé, Junghun, Masha, Clément, Marc, Maxime, Benjamin, Maria, Julien, Achraf, Naila, Théo, Gimo, Axel, Matías, Timur… Cet environnement unique, je le dois à Vianney et Patrick. Vous avez réussi à créer une équipe géniale qu'on retrouve rarement ailleurs. J'espère que les futurs membres perpétueront la tradition du respo gâteaux du mardi après-midi ! Merci aussi à Mélanie, pour ton soutien administratif sans faille, et sans qui tous ces voyages de dernière minute n'auraient pas été possibles.

Je remercie également tous les chercheurs rencontrés lors de conférences ou dans les couloirs des universités, dont les conseils et les discussions m'ont été particulièrement précieux. Merci en particulier à Romain, Joon, Panayotis, Bruno, Laurent, Atulya, José, Claire, Nicolas, Raimundo, et bien d'autres encore. Merci aussi aux participants du séminaire parisien de théorie des jeux et d'optimisation.

Merci à l'ENSAE et à tous les membres du CREST avec qui j'ai eu la chance d'échanger durant ces années : Jaouad, Zong, Nayel, Hugo, Anna, Matthieu, Arnak, Vincent, Victor-Emmanuel, Cristina, Austin, Nicolas, Kessang, Sirine, Clémentine, Redouane, Adrien.

Je n'oublie pas mes amis, du collège, du lycée, des prépas, des Mines, d'hier et d'aujourd'hui. Vous êtes ma bouffée d'oxygène, toujours là pour me faire rire, me réconforter ou m'écouter me plaindre. Ces liens me sont chers, et sachez que vous pourrez toujours compter sur moi comme j'ai compté sur vous. Je m'excuse aussi auprès des inconnus qui ont eu le malheur de recevoir une explication improvisée sur les inégalités de prophète à deux heures du matin, je promets de faire des efforts !

Enfin, je remercie toute ma famille, qui m'a soutenu avec ardeur depuis mon enfance. Et merci plus particulièrement à mes parents, mon frère, ma sœur, mes deux mamies chéries, et mes deux grands-pères à qui je pense très fort. Vous avez toujours cru en moi, et j'espère vous avoir enfin rendus fiers. Je vous aime très fort.

Si j'oublie quelqu'un, ce n'est ni par manque de gratitude, ni par manque de place dans mon cœur, mais simplement à cause de mon étourderie. J'espère que vous m'en excuserez!

Enfin, ces remerciements ne sont pas une clôture, mais une promesse, celle d'entretenir des liens durables et de parcourir ensemble de nouveaux chemins.

Contents

1 Introduction

> *"That's not fair!" — "You say that so often, I wonder what your basis for comparison is?"*
>
> **— Labyrinth (1986)**

From education [GS62], healthcare [RSÜ04], housing [BM01a], labor markets [HM05], to online advertisement platforms [EOS07], the problem of allocating limited resources to multiple agents, who may have diverse and potentially conflicting interests, lies at the heart of many decision-making systems.

Traditionally, the design of allocation rules has been driven by economic principles such as maximizing social welfare or achieving Pareto efficiency [Deb54]. However, in many settings, only focusing on utility can yield outcomes that are perceived as discriminatory or inequitable. This has led to a growing interest in integrating other ethical considerations, in particular fairness, as an additional constraint or alternate objective, directly into the algorithmic design of allocations.

In modern algorithmic settings, such challenges become even more pressing. Decisions that were once made by human judges, committees, or central planners are now increasingly delegated to automated systems. Indeed, advances in computing power and machine learning have enabled the widespread use of algorithms in high-stakes domains such as healthcare (e.g., tumor detection through deep learning [LKB+17]), finance (e.g., credit scoring using black-box models [Büc+22]), and employment (e.g., AI bots for interviewing applicants [OCo25]). While their efficiency and scalability make them attractive for broad deployment, they also introduce serious risks related to fairness and accountability [Pas15; Coo+22].

These concerns have been exacerbated by several high-profile failures. The COMPAS algorithm used in U.S. courtrooms was shown to exhibit racial bias in recidivism predictions [Lar+16]; the Gender Shades study revealed stark performance disparities in commercial facial recognition systems based on gender and skin tone [BG18a]; and Amazon's resume screening tool was found to systematically penalize female candidates [Das18]. These cases illustrate the risks of optimizing for performance without taking into account fairness constraints, at the risk of perpetuating or even amplifying existing biases.

Therefore, there is a growing need to design allocations that are sensitive to equity considerations [MS22], but doing so is far from straightforward. In dynamic settings, myopic or greedy approaches that optimize short-term utility can conflict with equity goals, which are long-term priorities. Satisfying these new constraints may also require the algorithm to explicitly use group-membership data to ensure parity in the allocation. And in strategic environments, relevant information needed to guarantee fairness may be private and must be carefully elicited through allocation mechanisms. New efficient algorithms must be designed

This thesis investigates how incorporating fairness affects the structure, analysis, and design of allocation algorithms.

1.1 What is Fairness?

In order to design and study optimal equitable allocations, we must first formalize what fairness is, and second how to frame equity in allocations problems. The idea of fairness plays a central role in how societies organize themselves, influencing how communities, institutions, and individuals interact with one another. It is a core part of our legal systems, but it is also a deeply moral and ethical notion: we are taught from an early age to learn to distinguish between which actions and outcomes should be perceived as fair and which as unfair. People have an intuitive understanding of whether they were wronged through a contract, or whether they received the right amounts of goods.

Yet, despite its seemingly strong natural aspect, providing an exact definition of fairness is an arduous task. What is deemed fair by an individual may be considered unfair by another, depending on cultural aspects, socio-economic context, or even personal beliefs. Even from a linguistic perspective, the notion of fairness is multifaceted. The English term fairness translates into French as justice, équité, impartialité, or honnêteté, each reflecting a distinct ethical principle. Dictionary definitions vary as well, some emphasize equality, others impartiality, reasonableness, or following established rules. Yet none of these definitions yields a single, operational definition. [1]

Although it is difficult to provide an exact definition, a recurring theme is that fairness is closely tied to the broader notion of justice. In particular, distributive justice [Ari99] is concerned with how resources should be allocated among individuals, directly connecting to the design of fair allocation rules. While other types of justice, such as procedural justice [Tyl90], emphasize the fairness of decision processes, our focus will be on distributive justice.

However, even within the scope of distributive justice, what an equitable allocation precisely means can remain unclear. Consider, for example, the task of assigning refugees to resettlement cities. What does it mean for a matching between migrants and cities to be just? Should fairness be defined in terms of equal numbers of people assigned to each city, or in terms of demographic parity, such as equal proportions of men and women? Should an equitable allocation prioritize migrants or emphasize city-level fairness? Should it focus on individuals, ensuring that each individual has a similar probability of being assigned to their top choice, or instead on group-level outcomes? Even for this seemingly simple application, multiple stakeholders may advocate for different fairness criteria, reflecting divergent objectives and priorities.

Furthermore, some fairness goals may be fundamentally unachievable due to incomplete or noisy information. Preferences can be stochastic and partially observable, or can be strategically hidden by self interested agents. Overall, each fairness definition, when applied to a specific allocation problem, creates distinct challenges arising from the interplay between feasible outcomes, available information, and additional equity requirements.

But adding additional constraints is not without negative impact on decision-making. It may restrict the set of feasible allocations, increase the amount of data necessary to achieve optimal fair outcomes, affect regret rates of online algorithms, or even introduce unavoidable performance trade-offs. Understanding how fairness can be embedded in complex decision-making processes, and how doing so significantly alters outcomes, is essential, both to improve algorithmic practices and to inform public policy.

[1] Reflecting this ambiguity, at the time of writing, the Wikipedia page for 'fairness" redirects to a disambiguation page (`https://en.wikipedia.org/wiki/Fairness`).

Our objective is not to define or find the "true" or "best" notion of fairness for allocations problems, a rather bold claim based on some of the discussion above, but rather starting from fairness notions that are mathematically well-defined, aligned with relevant ethical intuitions, and meaningful in specific decision-making contexts, we analyze their algorithmic consequences: How do fairness constraints alter the structure of feasible solutions? What trade-offs do they introduce? Under what conditions can fairness be achieved without sacrificing efficiency? What new algorithmic techniques are required? Another more informal title of this thesis could have been "How much and in what way does fairness gets in the way of traditional algorithm design and performance for a range of allocation problems?".

1.2 Outline and Contributions

This thesis is organized around 6 chapters, one chapter introducing general fairness definitions for allocation problems (Chapter 2), followed by five independent chapters, each presenting a self-contained research article. While each of those 5 chapters focuses on a distinct allocation problem and fairness notion, they collectively examine how fairness can be meaningfully incorporated into algorithmic decision-making, and to what extent this addition affects outcomes and classical algorithm design.

1. **Online Learning**

    - **Bandit Algorithms with Fairness Penalties**, Chapter 3: In online advertisement, sensitive attributes are becoming increasingly harder to access due to new regulations [Gua23], even though such information is critical to ensure equitable ad exposure across protected groups. We model a setting where an advertiser show ads and earns revenue from users arriving sequentially, but pays a fine based on the final allocation imbalance. To circumvent the non-availability of sensitive attributes, we consider that additional information regarding the agent, which helps to estimate these missing features, can be purchased at a price. We leverage techniques from the bandits literature and from online convex optimization to construct a primal dual online algorithm that evaluates the quality of the data on the fly, and achieves low regret.

      Based on "Trading-off Price for Data Quality to Achieve Fair Online Allocation", Mathieu Molina, Nicolas Gast, Patrick Loiseau, and Vianney Perchet, NeurIPS, 2023.

- **Bandits Algorithms with Soft Fairness Constraints**, Chapter 4: We consider a music streaming platform, which has to guarantee a certain level of revenue in the short term to artists that can be recommended. We model this problem as a classical stochastic bandit problem, but with additional short-term covering constraints, which ensure that each arm earns enough revenue. We introduce a family of algorithms that optimally balance the trade-off between constraint violation and learning regret.

  Based on "Multi-Armed Bandits with Guaranteed Revenue per Arm", Dorian Baudry, Nadav Merlis, Mathieu Molina Hugo Richard, and Vianney Perchet, AISTATS, 2024.

2. **Competitive Analysis**

- **Prophet Inequalities with Weaker Benchmark**, Chapter 5: We analyze a variant of the classical independent and identically distributed (i.i.d.) prophet inequality problem, which compares the performance of the best online policy to an offline benchmark, the "prophet". By weakening the offline benchmark, we show that the relative error rate decreases exponentially fast as the benchmark grows weaker. In Section 5.2, we further connect this relaxed benchmark to a group fair objective, and show how previous results can be directly leveraged for this fair allocation problem.

  Based on "Prophet Inequalities: Competing with the Top ℓ Items is Easy", Mathieu Molina, Nicolas Gast, Patrick Loiseau, and Vianney Perchet, SODA, 2025.

3. **Mechanism Design**

- **Efficient Equitable Auctions**, Chapter 6: We study auctions for allocating k identical goods among n symmetric unit-demand bidders. Uniform-price and discriminatory-price auctions are commonly used in practice, and are both efficient and revenue-equivalent under specific assumptions. We want to compare these two auctions along a third axis, equity, which represents the profit disparity among winners ex-post. We show that while uniform pricing can be more equitable in some cases, discriminatory pricing is strictly more equitable for log-concave signal distributions. Finally, we compute incentive compatible payment rules that achieve perfect ex-post equity while preserving both efficiency and revenue, showing that, surprisingly, perfect equity can be achieved at no cost.

Based on "Equitable Pricing in Auctions", Simon Finster, Patrick Loiseau, Simon Mauras, Mathieu Molina, and Bary Pradelski, EC, 2025 (Alphabetical Ordering).

4. **Optimization Under Constraints**

   - **Worst-Case Price of Fairness**, Chapter 7: We consider a structured allocation problem, where the set of feasible allocations is a matroid. We evaluate the worst-case relative loss in social welfare resulting from imposing group fairness constraints akin to equality of opportunity. The problem reduces to allocation over a polymatroid, and we show that in the worst-case, the price of fairness remains bounded and scales linearly with the number of protected groups. We further compute the price of fairness under additional structural and stochastic assumptions, and show under which conditions the social welfare loss is low or non-existent.

     Based on "The Price of Opportunity Fairness in Matroid Allocation Problems", Rémi Castera, Felipe Garrido-Lucero, Simon Mauras, Mathieu Molina, Patrick Loiseau and Vianney Perchet, Working paper (Alphabetical Ordering).

The following two articles are not included in this thesis:

- "Bounding and Approximating Intersectional Fairness Through Marginal Fairness", Mathieu Molina and Patrick Loiseau, NeurIPS, 2022,
- "On Preemption and Learning in Stochastic Scheduling", Nadav Merlis, Hugo Richard, Flore Sentenac, Corentin Odic, Mathieu Molina, and Vianey Perchet, ICML, 2023.

The next chapter will give a brief overview on how fairness has been formally defined in economics, computer science, and operations research literature, going from more philosophical concepts to mathematically sound definitions. We will examine key families of fairness definitions, individual vs. group fairness, ex-ante vs. ex-post fairness, and utility-based vs. constraints. Although not all definitions will be used, understanding the subtle differences between these fairness notions will be useful in grasping the various potential consequences on algorithmic design.

2 Fairness Definitions in Allocations Problems

In this section, we formally introduce several fairness notions developed to encode moral and philosophical considerations into algorithmic allocation decisions. We introduce the allocation settings of the five papers of this thesis in order to best compare the various fairness definitions. While the following chapters adopt domain-specific notations, this section uses common notations for ease of interpretation.

A General Allocation Model. We consider a population of $n \in \mathbb{N}$ agents, indexed by $i \in [n]$. Each agent receives an allocation x_i, and the total allocation vector $x = (x_1, \ldots, x_n)$ must lie within the set of feasible allocations $\mathcal{X}$. The feasibility constraints can be due to limited resources or other structural requirements. From the allocation x_i, the agent (or a central planner) derives a utility $u_i(x_i)$. Group-level fairness can be treated as though there is a meta-agent that represents the agents grouped together.

We briefly outline the five allocation problems studied in this thesis. These problems encompass a broad spectrum of settings, ranging from online to offline decision-making, stochastic to adversarial inputs, fully observed to partially observable environments, and fairness notions defined at the individual as well as the group level. They also include both strategic and non-strategic (static) scenarios. Despite this broad range of settings, the basic allocation problems all fit within the model described above.

1. **Online Advertising with Group-Fairness Fines** (Chapter 3): A company seeks to allocate ads to T arriving users, each from one of n unknown protected groups. Group membership and utility are not directly observable but can be estimated through costly signals. The feasible set is $\mathcal{X} = \prod_i \{0, 1, \ldots, n_i\}$, where n_i is the number of users from group i. Fairness may require proportional exposure across groups, i.e., $x_i / n_i = x_j / n_j$.

2. **Music Streaming with Revenue Guarantees** (Chapter 4): A music streaming platform must recommend songs to a stream of T users, selecting among n different artists. When artist i is recommended, they receive an i.i.d. reward

U_i with unknown mean μ_i. The platform seeks to maximize utility while ensuring that each artist i earns a minimum expected revenue of $\lambda_i T$. Notably, the revenue guarantees must be satisfied in expectation even over short time horizons. The set of feasible allocations is $\mathcal{X} = \{x \in \mathbb{R}^n_+ : \sum_i x_i \leqslant T\}$, and each artist's expected utility is given by $u_i(x_i) = \mu_i x_i$.

3. **Hiring with Diversity Incentives** (Chapter 5): A firm must hire **2** individuals from a sequence of T arriving applicants, each randomly assigned to one of two protected groups (e.g., men and women). The feasible allocations are subsets $x \subseteq A_1 \times A_2$ with $|x| \leqslant 2$, where A_1 and A_2 denote the (random) partitions of the applicants into the two groups. The utility is the sum of the values of the selected candidates; however, a penalty is incurred when both selected individuals belong to the same group, so as to incentivize diversity of selected applicants.

4. **Reducing Profit Disparity when Selling Goods** (Chapter 6): An auctioneer sells k identical items to n unit-demand buyers. The allocation space is $\mathcal{X} = \{x \in \{0,1\}^n : \sum_i x_i \leqslant k\}$. Buyers have private or common valuations v_i, and utility is $u_i = v_i - p_i$. The objective is to design mechanisms that ensure efficiency and reduce utility dispersion among winners, to fairly share the realized total profit, even though the utility is not known to the auctioneer.

5. **Social Welfare Loss of Fair Refugees Resettlement** (Chapter 7): A central planner allocates places in different cities for refugee resettlement under matroid constraints, across n demographic groups. Utility is cardinal, $u_i(x_i) = x_i$, and the feasible allocations set $\mathcal{X}$ is a polymatroid of dimension n. Fairness requires that each group receive the same proportion of the maximum allocation it could achieve in isolation.

Having introduced the relevant allocation settings, we now survey the principal fairness notions developed across theoretical economics, operations research, computer science, and statistics. These definitions are organized thematically, and their presentation follows a roughly chronological progression, reflecting how fairness concepts have evolved across disciplines. We identify four broad categories of fairness notions:

- **Fair Division:** Concerned with the allocation of divisible or indivisible goods among agents, using notions such as proportionality, equitability, and envy-freeness, defined with respect to agents' subjective valuations.

- **Game Theory:** Focuses on the axiomatic foundations of fair surplus or utility allocation in transferable utility cooperative games, typically through bargaining solutions (e.g., Nash bargaining) or the Shapley value.
- **Operations Research:** Adopts a quantitative perspective on fairness, emphasizing different social welfare objectives and fairness measures such as max-min fairness, α-fairness, and Jain's index.
- **Machine Learning:** Rooted in a statistical perspective on decision-making, encompassing both group-level criteria (e.g., demographic parity, equal opportunity), and individual-level notions of fairness, particularly in settings with uncertainty.

While not all of the fairness definitions discussed in this section will be studied explicitly in the subsequent chapters, this overview aims to convey the richness and diversity of fairness concepts and to illustrate how these notions interact with different allocation settings and impact algorithm design. The remainder of this chapter develops each family of fairness notions in greater detail and identifies the specific definitions applied in the five allocation problems introduced earlier.

2.1 Fair Division and Preference-Based Fairness

One of the earliest and most influential domains in formalizing fairness is *fair division*, which addresses how to allocate a limited amount of heterogeneous goods among agents such that each agent subjectively perceives their share to be fair. A central feature of this literature is the distinction between *divisible* goods (e.g., a cake) and *indivisible* goods (e.g., gifts). The original example, often attributed to Steinhaus, Knaster, and Banach in the 1940s [Ste48], is the *cake-cutting* problem: how to divide a cake among multiple agents with subjective valuations.

There are three main fairness concepts that arise from fair division: equitability, proportionality (or more generally, fair-share), and envy-freeness. See [Mou03; Bra+16] for comprehensive overviews of fair division, and more broadly computational social choice.

2.1.1 Equitability

Equitability demands that all agents derive the same subjective value from their allocated shares [Fre+19].

Definition 2.1. An allocation $x \in \mathcal{X}$ is said to be *equitable* if $u_i(x_i) = u_j(x_j)$ for all $i, j \in [n]$.

This notion reflects an ideal of fairness from the standpoint of equal satisfaction. However, equitability does not ensure that each agent receives a sufficiently large share, only that everyone derives the same utility from the allocation, even if that utility is low.

2.1.2 Proportional Fairness (Fair-Share)

In contrast to equitability, proportional fairness requires that each agent receives a "big enough" share of the total value of the goods.

Definition 2.2. An allocation $x \in \mathcal{X}$ is *proportionally fair* if for all $i \in [n]$, $u_i(x_i) \geqslant \max_{x \in \mathcal{X}} u_i(x_i)/n$.

This ensures that each agent receives at least a $1/n$ share of their total possible utility. While such allocations are guaranteed to exist for divisible goods under mild conditions [DS61], they may not exist in the case of indivisible goods.

More recently, [BF22] introduced the notion of *fair share functions*, generalizing proportional fairness. For each utility function u_i and number of agents n, we define a fair share $s(u_i, n)$, and an allocation is considered fair if $u_i(x_i) \geqslant s(u_i, n)$. This notion encompasses both proportional fairness and other criteria like the maximin share.

Problem 4 somewhat falls into this category: each of the n artists must receive at least a minimum expected utility of $\lambda_i \cdot T$, which corresponds to the fair-share constraint $u_i(x_i) \geqslant \lambda_i T$.

2.1.3 Envy-Freeness

While proportionality focuses on agents' individual satisfaction, envy-freeness takes into account comparisons between agents. Introduced by Gamow and Stern [GS58], an envy-free allocation ensures that no agent prefers another agent's allocation to their own.

Definition 2.3. An allocation $x \in \mathcal{X}$ is *envy-free* if for all $i, j \in [n]$, $u_i(x_i) \geqslant u_i(x_j)$.

This condition prevents envy by guaranteeing that each agent finds their own bundle at least as desirable as anyone else's. For divisible goods with continuous, additive preferences, envy-free allocations are known to exist. However, for indivisible goods, envy-free allocations often do not exist; for example, when one item must be assigned to one of two agents, envy is unavoidable.

Given the nonexistence of envy-free or proportionally fair allocations in settings with indivisible goods, several relaxations have been proposed.

One common relaxation is *envy-freeness up to one good* (EF1), which requires that any envy can be eliminated by removing a single item from the envied bundle. This notion can be satisfied more easily than envy-freeness, for example an EF1 allocation always exists under weakly additive valuations [Car+16].

2.1.4 Ex-ante vs Ex-post Fairness of Randomized Allocations

An alternative way to circumvent impossibility results is to allow randomized allocations. A *random allocation* is a probability distribution over feasible deterministic allocations. Ex-ante fairness ensures that fairness conditions are satisfied in expectation.

Definition 2.4. A randomized allocation $X = (X_1, \dots, X_n)$ satisfies *ex-ante proportional fairness* if for all $i \in [n]$, $\mathbb{E}[u_i(X_i)] \geqslant \max_{x \in \mathcal{X}} u_i(x_i)/n$.

Such allocations always exist, e.g., by assigning the entire bundle to agent i with probability $1/n$ for each i. More refined mechanisms, such as the *Probabilistic Serial* rule [BM01a], achieve envy-freeness and proportionality ex-ante under certain conditions. The resulting fractional allocation can often be implemented as a lottery over integral allocations using the Birkhoff-von Neumann decomposition of bi-stochastic matrices.

In problems 4 and 7, we rely on randomized (or fractional) allocations to satisfy fairness constraints while maintaining feasibility. In particular, in Problem 7, we further show how an ex-ante optimal allocation can be transformed into an approximately fair ex-post allocation. Taking an ideal allocation ex-ante, while retaining good properties ex-post, is a common "best of both worlds" consideration prevalent in fair division.

Equitability, proportionality, and envy-freeness reflect distinct fairness ideals. None implies the others in general, and they can conflict. While envy-freeness compares the value of different pieces to the same person, equitability compares the value of different pieces to different agents. In fact, these three notions often conflict, and in general, no allocation may satisfy two of them simultaneously. This type of tension between different fairness notions is ubiquitous in the fair allocation literature, even beyond fair division. As a result, most works focus on a single fairness definition, with its pros and cons.

But while these different fairness notions can each be seen as reasonable candidates for defining fair allocations, they can still conflict with one another. This naturally raises the following question: do these mathematical definitions truly capture the ethical ideas they are meant to reflect? The next section turns to fairness notions rooted in game-theoretic models, particularly cooperative games. Here, fairness is defined axiomatically: one begins by stating the properties that a fair allocation *should* satisfy, and then derives the allocation implied by these axioms, if this allocation exists.

2.2 Cooperative Game Theory and Axiomatic Fairness

Fairness has also been a central concern in game theory, particularly in cooperative game theory, where the focus is on how to fairly distribute a collective payoff among a group of agents who can choose to participate in a coalition. In contrast to fair division, this domain explicitly considers incentives: agents collaborate to generate value, and a fair outcome must reward each participant proportionally to their contribution. This perspective gave rise to two landmark fairness concepts: the *Nash bargaining solution* and the *Shapley value*. For a general treatment of cooperative games, see [Mye91], and for a fairness-focused perspective, see [You94].

As mentioned above, the distinctive feature of this domain is its *normative* approach to fairness. The existence and uniqueness of a fair allocations follows from the axioms that such a solution should satisfy.

2.2.1 Bargaining Solutions

John Nash proposed an axiomatic solution for two-player bargaining problems in 1950 [Nas50], later extended to multiple players. Consider two agents negotiating over a feasible set $\mathcal{X}$ of utility pairs $(u_1(x_1), u_2(x_2))$, with $(u_1(d_1), u_2(d_2))$ denoting their disagreement utilities. That is, they either agree on a feasible outcome $x \in \mathcal{X}$, or receive default $(u_1(d_1), u_2(d_2))$ if negotiations fail. A fair solution must satisfy:

1. **Pareto Efficiency:** The allocation is not strictly dominated by another.
2. **Symmetry:** If the feasible utilities are symmetric, then players should receive equal utilities.
3. **Invariance to Affine Transformations:** Fairness is preserved under rescaling of utilities.
4. **Independence of Irrelevant Alternatives:** Removing unchosen allocations from the feasible set should not change the fair allocation.

Under these axioms, Nash proved the unique fair outcome is the allocation maximizing the product of marginal gains: $x^N = \arg\max_{x \in \mathcal{X}} (u_1(x_1) - u_1(d_1)) \cdot (u_2(x_2) - u_2(d_2))$.

An alternative is the *Kalai-Smorodinsky (KS) solution* [KS75], which replaces Independence of Irrelevant Alternatives with:

5. **Monotonicity:** If one agent's achievable utility improves while the other's does not worsen, then the fair allocation should also have an improved utility.

Geometrically, when the utility frontier is convex, the KS solution lies at the intersection of the Pareto frontier with the line connecting the disagreement point $(u_1(d_1), u_2(d_2))$ to the point of maximum utilities for both agents $(\max_x u_1(x_1), \max_x u_2(x_2))$. Unlike Nash, the KS solution equalizes relative gains:

$$\frac{u_1(x_1^{KS}) - u_1(d_1)}{\max_x u_1(x_1) - u_1(d_1)} = \frac{u_2(x_2^{KS}) - u_2(d_2)}{\max_x u_2(x_2) - u_2(d_2)}.$$

This formulation prioritizes agents with higher potential gain, taking into account that agents can have different entitlements depending on the specific allocation problem, where for KS the entitlement is proportional to the best achievable utility in isolation.

In Chapter 7, the group fairness objective can be interpreted as the geometric (not axiomatic) generalization of the KS solution to multiple agents. The fairness constraint aims to equalize scaled utilities across groups, where scaling is done with respect to the maximum utility a group could receive if treated alone, which corresponds exactly to the equalization of relative gains of the KS solution. This can also be interpreted as a form of weighted equitability. However, unlike KS, the allocation is not necessarily Pareto efficient, highlighting the efficiency loss as the number of groups grows beyond 2.

2.2.2 The Shapley Value

Bargaining solutions focus on joint decision-making, but in games with more than two players, some subgroups may prefer to form separate coalitions. To address such scenarios, cooperative game theory introduces the *core* [Gil59; Aum60], a stability concept. A cooperative game with n agents is defined via a *characteristic function* $v : 2^{[n]} \to \mathbb{R}$, where $v(S)$ denotes the value that coalition S can generate. The game has *transferable utility*, meaning the total utility $v(S)$ can be redistributed within a coalition.

Definition 2.5. An allocation $x \in \mathbb{R}^n$ is *in the core* if for all $S \subseteq [n]$, we have $\sum_{i \in S} x_i \geqslant v(S)$.

This stability notion says that if players do not receive enough surplus, they may deviate, form a coalition, and not participate in the game.

The *Shapley value* introduced by Shapley in 1951 [Sha53], is a canonical way to share surplus in a transferable utility game. It is defined as satisfying four axioms: Efficiency (all the value $v([n])$ is distributed), Symmetry, Linearity, and not giving surplus to players who do not contribute to anything.

Shapley proved that there is a unique value function φ_i, where φ_i is the utility that should be distributed to agent i, that satisfies all four axioms:

$$\varphi_i(v) = \sum_{S \subset [n] \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} (v(S \cup \{i\}) - v(S)),$$

that is to say, φ_i is the expected marginal contribution of agent i over all possible joining orders. While the core may be empty in general, a key result is that for convex games, those where v is supermodular, the Shapley value lies in the core [Sha71], thus satisfying both fairness and stability.

The Shapley value has seen broad applications, from cost-sharing and voting power analysis to machine learning interpretability [Roz+22]. While very useful, it assumes full knowledge of $v(S)$ for all coalitions S and can also be computationally hard (computing the Shapley value is NP complete [DP94]).

2.2.3 Fairness and Strategic Manipulations

One important connection between fair allocations and game theory lies in the tension between fairness goals and strategic behavior or stability requirements. In school choice, for example, it has been shown that stable matchings may be incompatible with fairness constraints such as diversity quotas across student types [Koj12]. Similarly, colleges cannot guarantee equal hiring opportunities across student types when faced with rational employers [KRZ18].

In Chapter 6, incentive considerations are central: agents have private valuations, and the goal is to design mechanisms that not only allocate goods efficiently but also fairly share surplus in a way that encourages truthful reporting. Fairness solutions like the Shapley value inform how profits should be redistributed to promote participation.

While the normative approach to fairness provides a powerful justification for adoption, it typically yields a coarse binary classification: an allocation either satisfies the fairness axioms or it does not. Indeed, this does not take into account cases where no exact fair allocation exists or when trade-offs are necessary. In such cases, it becomes essential to adopt a more quantitative approach to fairness, by defining different fairer objectives and fairness metrics, making it possible to compare the fairness of different allocations.

2.3 Operations Research and Quantitative Fairness

2.3.1 A Shift in the Social Objective

In 1971, the philosopher John Rawls published *A Theory of Justice* [RAW71], a work that has profoundly influenced formal fairness concepts in economics and social sciences, remaining one of the most important theories on fairness to this day. Rawls proposes a thought experiment in which individuals design the rules of society behind a *veil of ignorance*, that is to say without knowning their own position within it. In our context, this translates to choosing allocation rules without knowing which which agent we will become. From this perspective, rational agents would favor rules that protect the worst-off, since they could end up in that position.

Mathematically, this leads to maximizing the utility of the worst-off agent [Sen17]: this is max min fairness.

Definition 2.6. An allocation is *egalitarian*, or *max-min fair*, if it solves the optimization problem $\max_{x \in \mathcal{X}} \min_{i \in [n]} u_i(x_i)$.

While max-min fairness is egalitarian, it can lead to an important total efficiency loss. For example, it treats utilities $(10, 10, 10)$ and $(10, 100, 100)$ as equally fair, since the minimum utility is 10 in both, whereas the total utility is quite different. To refine this, *leximin fairness* sequentially maximizes the utility of the worst-off, then the second-worst, and so on [Ham76; DG77]. Leximin allocations are always Pareto efficient.

One crucial difference from the previously discussed notions is that here, fairness is embedded directly into the optimization objective: rather than requiring the allocation rule to satisfy specific properties, we redefine the social welfare function itself, changing from a utilitarian objective to an egalitarian one.

2.3.2 General Fair Objectives

With the growth of computer networks in the 1980s and 1990s, fair allocation of bandwidth and other shared resources became central. Max-min and leximin fairness were adapted to queuing systems [Nag85] and network resource sharing [BG86]. A broader class of fairness objectives, known as α-*fairness*, was introduced in [MW98]:

Definition 2.7. For $\alpha \geqslant 0$, an allocation x is α-fair if it maximizes

$$S_\alpha(x) = \begin{cases} \sum_{i \in [n]} \frac{u_i(x_i)^{1-\alpha} - 1}{1-\alpha}, & \text{if } \alpha \neq 1, \\ \sum_{i \in [n]} \log(u_i(x_i)), & \text{if } \alpha = 1. \end{cases} \tag{2.1}$$

This objective is continuous in α. When $\alpha = 0$, it reduces to the utilitarian welfare; for $\alpha = 1$, it corresponds to the Nash bargaining solution, as maximizing the sum of logarithms is equivalent to maximizing the product of utilities; and as $\alpha \to \infty$, it converges to the max-min fair allocation. By tuning α, we can trade off between efficiency and equity. This illustrates how a carefully designed social welfare function can recover some normative fairness properties. See [CH23] for formulating fairness as an optimization problem.

2.3.3 Measuring Fairness

In many applications, rather than changing the optimization objective entirely, we would like to balance the standard utilitarian goals with fairness. This leads to *fairness metrics* that quantify how fair an allocation is. Well-known examples include the *Gini coefficient* [Gin12a]:

$$\text{Gini}(x) = \frac{\sum_i \sum_j |u_i(x_i) - u_j(x_j)|}{2n \sum_i u_i(x_i)}, \tag{2.2}$$

and *Jain's fairness index* [JCH98]:

$$\text{Jain}(x) = \frac{\left(\sum_i u_i(x_i)\right)^2}{n \sum_i u_i(x_i)^2} = \frac{1}{1+c}, \tag{2.3}$$

where c is the coefficient of variation. Both metrics equal 1 when all utilities are equal, and are respectively equal to 0 and $1/n$ in the worst-case.

The main advantage of such metrics is that they induce a total order on allocations (through the usual order on $\mathbb{R}$). This allows direct comparison of allocations in terms of fairness, and we can treat the fair allocation problem as a multi-objective optimization problem in terms of total utility and fairness.

In Chapter 6, we take exactly this approach: we define a fairness metric that measures the dispersion of utilities among winners in an auction, allowing us to assess and compare the fairness of different auction formats.

2.3.4 Desirable Properties of Social Welfare Functions and Fairness Measures

More generally, what makes an objective or measure "good" with respect to fairness? One common desirable criterion is the *Pigou-Dalton principle* [Pig12a; Dal20a], which states that transferring utility from a better-off to a worse-off agent should improve the fairness score.

For symmetric functions, this principle is equivalent to *strict Schur-concavity*. All objectives and measures discussed above, α-fairness (for $\alpha > 0$), the Gini coefficient, Jain's index, and our proposed fairness measure in Chapter 6, satisfy this property.

Yet, while quantitative fairness provides a flexible way to balance efficiency and equity in allocation problems, it often assumes that the inputs and preferences are given and reliable. In many modern applications, however, decisions are based on large-scale datasets that may themselves reflect historical and systemic biases. This has shifted attention toward understanding how automated decisions can sustain or even amplify group-based discrimination. As a result, the machine learning community has developed a rich literature on fairness under uncertainty, focusing on statistical definitions of fairness that can detect and mitigate such biases across groups and individuals.

2.4 Machine Learning and Statistical Fairness

While fairness definitions in machine learning typically come from classification or regression tasks, rather than allocation problems, this recent literature has significantly influenced how fairness is conceptualized in allocation. In particular, it has brought group-level fairness considerations to allocation settings.

The rise of social justice movements, especially since the mid-20$^{\text{th}}$ century, has emphasized the importance of accounting for systemic bias and discrimination. In response, important anti-discrimination laws were enacted, such as the U.S. Equal Employment Opportunity Commission's "four-fifths rule" (or 80% rule) in the 1970s, which required hiring rates across demographic groups to meet minimal proportional thresholds. These developments laid the groundwork for early notions of statistical fairness in testing and decision-making, beginning in the 1970s [Cle66; Dar71]. The field has grown sharply since the 2010s, due to the integration of machine learning algorithms in high-stakes decision processes. See [BHN23] for a comprehensive treatment of fairness in machine learning.

2.4.1 Statistical Group Fairness

We begin with a binary classification setting to introduce statistical fairness notions. Each individual i is represented by an i.i.d. sample $(Z_i, A_i, Y_i) \in \mathbb{R}^d \times \mathbb{R}^p \times 0, 1$ drawn from a joint distribution, meaning that Z_i, Y_i, and A_i can be dependent. Here, Z are features (e.g., education, location), A are sensitive attributes (e.g., gender, ethnicity), and Y is the binary outcome to be predicted (e.g., loan approval). The goal is to learn a classifier $h : \mathbb{R}^{p+d} \to 0, 1$, where $h(Z, A)$ predicts Y based on the input features.

Standard learning focuses on minimizing prediction error, for instance $\mathbb{E}[\mathbb{1}[h(Z, A) \neq Y]]$. However, because this objective is the average over the entire population, it may hide significant differences in error rate across groups, measured by $\mathbb{E}[\mathbb{1}[h(Z,A) \neq Y] \mid A]$. Moreover, simply omitting A from the input does not eliminate discrimination due to correlations between Z and A. Hence, explicit fairness definitions are needed.

Many group-fairness notions can be formalized through independence or conditional independence between true labels, predicted labels, and sensitive features:

Definition 2.8. A classifier h satisfies

- Demographic Parity if $A \perp\!\!\!\perp h(Z,A)$,
- Equality of Opportunity if $(A \perp\!\!\!\perp h(Z,A)) \mid Y = 1$,
- Predictive Parity if $(A \perp\!\!\!\perp Y) \mid h(Z,A) = 1$.

One of the earliest and simplest definitions is *demographic parity*, which requires that prediction outcomes be independent of the sensitive attributes [CV10]. This notion, also called *statistical parity*, can be problematic when A is correlated with the true label Y, leading to a potentially important loss in accuracy. For example, if age correlates with job performance in a physically demanding role, enforcing demographic parity might force a company to ignore important predictors to determine the suitability of job applicants. To address this, *equality of opportunity* [Har+16b] requires equal true positive rates across groups. Yet another notion, *predictive parity* [Cho16], conditions on the predictor instead of the true label.

These classification-based definitions translate naturally to allocation settings. For instance, suppose that $x_i \in [0,1]$ denotes the probability that agent i is selected (or allocated a resource). Then, fairness for the allocation corresponds to fairness in the empirical distribution derived from the allocation. For example, demographic parity in this context requires that the average allocation is equal across all protected groups:

$$\frac{\sum_{i:A_i=a} x_i}{|\{i : A_i = a\}|} = \frac{\sum_{i:A_i=a'} x_i}{|\{i : A_i = a'\}|}, \forall a, a' \in \mathrm{Im}(A). \tag{2.4}$$

As in fair division, there are multiple impossibility results showing that these three fairness definitions are incompatible in general [BHN23]. Furthermore, when A has multiple components (e.g., ethinicity and gender), marginal group fairness does not imply *intersectional fairness*, fairness across all combinations. This leads to statistical challenges due to the exponential number of subgroups [Kea+17].

Fairness notions in Problems 3, 5, and 7 are grounded in machine learning fairness notions, focusing on group-based fairness with respect to sensitive attributes. However, when A is not directly observed, as in Problem 3, the uncertainty about group membership impacts both the utility and fairness of the allocation. Furthermore, in Problems 3 and 5, we do not require the allocation to be group-fair, but rather impose penalties depending on how unfair the allocation is.

2.4.2 Achieving Fairness

While a constant classifier always achieves demographic parity and equality of opportunity, it is not always clear how to achieve a fair classifier with non-trivial performance. There are three types of general methods:

- **Pre-processing:** transform the data to remove correlations between Z and A,
- **In-processing:** optimize the classifier under fairness constraints,
- **Post-processing:** modify outputs (e.g. randomization) of an existing classifier to satisfy fairness constraints.

When fairness constraints are too strict, they can be relaxed. For example, demographic parity can be relaxed via metrics:

$$|\mathbb{P}(h(Z, A) = 1 \mid A = a) - \mathbb{P}(h(Z, A) = 1 \mid A = a')| \leqslant \varepsilon, \tag{2.5}$$

or using mutual information:

$$\mathrm{MI}(h(Z, A); A) \leqslant \varepsilon, \tag{2.6}$$

which reduces to 0 when independence holds.

2.4.3 Individual Fairness

While these definitions enforce some level of fairness across groups, discrimination might still happen at the individual level. To address this, [Dwo+11] introduced the notion of *individual fairness*, which seeks to ensure that similar individuals are treated similarly. Let $H(Z, A) \in [0, 1]$ be a randomized classifier, D be a metric on the outcome space (here $[0, 1]$), and d be a similarity metric between agents. Then:

Definition 2.9. A randomized classifier H is individually fair with respect to (d, D) if

$$D(H(Z_i, A_i), H(Z_j, A_j)) \leqslant d(i, j), \quad \forall i, j \in [n]. \tag{2.7}$$

This definition can be interpreted as a Lipschitz constraint on the randomized classifier. The main challenge with this approach lies in defining a meaningful similarity metric between agents. In practice, such a metric is rarely objective and may itself encode societal biases.

In Problem 6, we explore a fairness notion that can be related to individual fairness: we show that certain auction formats yield utility across winners that are Lipschitz-continuous in the agents' private signals. This limit the dispersion of winners' utility, ensuring that the total profit is shared more evenly among agents which do receive an item.

The fairness notions introduced so far span a wide spectrum, and we have highlighted how these distinct definitions interact with structural constraints, incentives, uncertainty and algorithmic design in diverse ways.

The remainder of this thesis is structured into five chapters, each developing theoretical results for a specific combination of an allocation problem and a corresponding fairness definition.

3 Trading-off data quality to achieve fair online allocation

Abstract. We consider the problem of online allocation subject to a long-term fairness penalty. Contrary to existing works, however, we do not assume that the decision-maker observes the protected attributes—which is often unrealistic in practice. Instead they can purchase data that help estimate them from sources of different quality; and hence reduce the fairness penalty at some cost. We model this problem as a multi-armed bandit problem where each arm corresponds to the choice of a data source, coupled with the online allocation problem. We propose an algorithm that jointly solves both problems and show that it has a regret bounded by $\mathcal{O}(\sqrt{T})$. A key difficulty is that the rewards received by selecting a source are correlated by the fairness penalty, which leads to a need for randomization (despite a stochastic setting). Our algorithm takes into account contextual information available before the source selection, and can adapt to many different fairness notions. We also show that in some instances, the estimates used can be learned on the fly.

Contents

3.1 Introduction

We consider the problem of online allocation with a long-term fairness penalty: A decision maker interacts sequentially with different types of users with the global objective of maximizing some cumulative rewards (e.g., number of views, clicks, or conversions); but she also has a second objective of taking globally "fair" decisions with respect to some protected/sensitive attributes such as gender or ethnicity (the exact concepts of fairness will be discussed later). This problem is important because it models a large number of practical online allocation situations where the additional fairness constraint might be crucial. For instance, in online advertising (where the decision maker chooses to which users an ad is shown), some ads are actually positive opportunities such as job offers, and targeted advertising has been shown to be prone to discrimination [LT19b; Spe+18a; Ali+19a]. Those questions also arise in many different fields such as workforce hiring [Dic+18], recommendation systems [Bur17], or placement in refugee settlements [Aha+21].

There has been a flourishing trend of research addressing fairness constraints in machine learning—see e.g., [Cho17; HPS16; KR18; BHN19; Eme+20; ML22a]—and in sequential decision-making problems—see e.g., [Jos+16; Jab+17; HK18]—as algorithmic decisions have real consequences on the lives of individuals, with unfortunate observed discrimination [BG18a; Lar+16; Das18]. In online allocation problems, general long-term constraints have been studied for instance by [AD15] who maximize the utility of the allocation while ensuring the feasibility of the average allocation, or by [BLM20] who show how to handle hard budget constraints for online allocation problems. More directly related to fairness, [NT20] formulate the problem of fair repeated auctions with a hard constraint on the difference between the number of ads shown to each group. Finally, in recent works, [BLM21; Cel+22] consider the online allocation problem where a non-separable penalty related to fairness is suffered by the decision maker at the end of the decision-making process instead of a hard constraint—these works are the closest to ours.

All the aforementioned papers, unfortunately, assume that the protected attributes (which define the fairness constraints) are observed before taking decisions. In practice, it is often not the case—for instance to respect users' privacy [Gua23]—and this makes it challenging to satisfy fairness constraints [LMC18]. In online advertising for example, the decision-maker typically has access to some public "contexts" on each user, from which she could try to infer the value of the attribute; but it was shown that the amount of noise can be prohibitive and therefore ensuring that a campaign reaches a non-discriminatory audience is non-trivial [Gel+20].

Our contribution. In this paper, we consider the online allocation problem under long-term fairness penalty, in the practical case where the protected attributes are not observed. Instead, we consider the case where the decision-maker can pay to acquire more precise information on the attributes (beyond the public context), either by directly compensating the user (the more precise the information on the attribute, the higher the price) or by buying additional data to some third parties data-broker.[1] Using this extra information, she should be able to estimate more precisely, and thus sequentially reduces, the unfairness of her decisions while keeping a high cumulative net reward. The main question we aim at answering is *how should the decision maker decide when, and from which source (or at what level of precision), to buy additional data in order to make fair optimal allocations?*

Compared to the closest existing works [BLM21; Cel+22] which study online allocation with a long term fairness penalty, the main novelty in the setting we examine is two-fold: we allow for uncertainty on the attributes of each individual, and more importantly we consider *jointly* the fair online allocation problem with a source selection problem. Consequently, we present the efficient algorithm 1 that tackles both of these challenges concurrently. This algorithm combines a dual gradient descent for the fair allocation aspect and a bandit algorithm for the source selection part. The final performance of an algorithm is its net cumulative utility (rewards minus costs of buying extra information) penalized by its long-term unfairness; that is quantified by the "regret": the difference between this performance and the one of some benchmarking "optimal" algorithm. We show that algorithm 1 has a sub-linear regret bound under some stochastic assumptions. Notably, the performance achieved by algorithm 1 using randomized source selection is strictly better than when using a single fixed source, because of the interaction through the fairness penalty—a key difference with standard bandit settings. On a more technical level we show how one can model the randomness and estimates for the protected attributes, how to bound the fair dual parameters which is crucial in order to use adversarial bandit techniques, and how to combine the analysis of the primal and dual steps of the algorithm.

There are many different definitions of group fairness that can be studied (e.g., demographic parity, equal opportunity, etc.). Instead of focusing on a specific one, we consider a generic formulation that can be instantiated to handle most of those different concepts (see section 3.2.1 and section 3.5). We also discuss in section 3.4.1 how to adapt our algorithm to different fairness penalties. This gives a higher level

[1]This setting includes as special cases the extremes where no additional information can be bought and where the full information is available.

of generality to our results compared to existing approaches that can handle fewer fairness criteria (e.g., [Cel+22]).

For the sake of clarity, we expose our key results in a simple setting. In particular, we assume binary decisions and a linear utility, and we assume that the expected utility conditional on the context is known. All these assumptions can be relaxed. In particular, we can learn the utilities (see section 3.4.2). We can handle also more general decision variables and utility forms. This allows in particular to tackle problems of matching and auctions (albeit with some restrictions), we discuss that in sections 3.6 and 3.7.

Related work. The problem of online fair allocation is closely related to online optimization problems with constraints, which is studied in a few papers. For instance, bandit problems with knapsack constraints where the algorithm stops once the budget has been depleted have been studied by [BKS18a; AD16a]. [LY21; LSY22] consider online linear programs, with stochastic and adversarial inputs. [Liu+21c] deal with linear bandits and general anytime constraints, which can be instantiated as fairness constraints. More recently [CCK22; Cas+22a] propose online algorithms for long-term constraints in the stochastic and adversarial cases with bandit or full information feedback. Some papers take into account soft long-term constraints [AD15; Jen+16], and more recently in [BLM21; Cel+22] where the long-term constraint can be instantiated as a fairness penalty—we also adopt a soft constraint. We depart from this literature by considering the case where the protected attributes (based on which fairness is defined) is not observed. We consider the case where the decision-maker can buy additional information to estimate it (which adds considerable technical complexity), but even in the case where no additional information can be bought our work extends that of [BLM21; Cel+22].

As mentioned above, the fairness literature usually assumes perfect observation of the protected attributes yet noisy realizations of protected attributes or limited access to them to measure and compute fair machine learning models has also been considered [Lam+19; Cel+21; Zha+22]. In some cases, the noise may come from privacy requirements, and the interaction between those two notions has been studied [Jag+19; CS21], see [Fio+22] for a survey. There are also works on data acquisition, which is similar to purchasing information from different sources of information; e.g., [Che+18; CZ19] study mechanisms to acquire data so as to estimate some of the population's statistics. They use a mechanism design approach where the cost of data is unknown and do not consider fairness (or protected attributes). However, none of these approaches can handle sequential decision problems with fairness constraints or penalties (and choosing information sources).

3.2 Preliminaries

3.2.1 Model and assumptions

We present here a simpler model, and later-on discuss possible extensions. Consider a decision maker making sequential allocation decisions for a known number of T users (or simply stages) in order to maximize her cumulative rewards. The user t has some protected attributes $a_t \in \mathcal{A} \subset \mathbb{R}^d$, that is not observed before taking a decision x_t. On the other hand, the decision-maker first observes some public context $z_t \in \mathcal{Z}$, where $\mathcal{Z}$ is the finite set of all possible public contexts and she has the possibility to buy additional information. There are K different sources for additional information and choosing the source $k_t \in [K]$ has a cost of p_{k_t}, but it provides a new piece of information, which together with the public information is summarized in the random variable $c_{tk_t} = (z_t, \text{data from } k_t)$. Based on this, the decision $x_t \in \{0, 1\}$ can be made; this corresponds to include, or not, user t in the cohort (for instance, to display an ad or not). Including user t generates some reward/utility u_t, which might be unknown to the decision maker (as it may depend on the private attribute), but can be estimated using the different contexts.

To fix ideas, we show how this model applies to two examples: Imagine an advertiser aiming to display ads to an user, able to see some bare-bone information through cookies, such as which website was previously visited (z_t). Based on this information, they can decide whether or not to buy additional information (c_{tk}) from different data brokers that collect user activity. For example, if they observe that the user has browsed clothing stores, they might opt to acquire data containing purchase details from this website. This enables them to estimate the user's gender (a_t) based on the type of clothing bought. Now consider the problem of fairly relocating refugees to different cities. When the organization in charge of resettlement receives a resettlement case (z_t), it can either decide to directly assign the refugee to a specific city, or to conduct an additional costly investigation (which might involve a third party watch-dog) to get more information (c_{tk}) on some protected attributes of interest such as wealth or age (a_t), which might have been intentionally misreported.

We assume that the global objective is to maximize the sum of three terms: the cumulative rewards of all selected users, minus the costs of the additional information bought, minus some unfairness penalty, represented by some function $R(.)$. Denoting by $\boldsymbol{k} = (k_1, \cdots, k_T)$ and $\boldsymbol{x} = (x_1, \cdots, x_T)$ the sources and allocations selected during the T rounds, the total utility of the decision maker is then

$$\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) = \sum_{t=1}^{T} u_t x_t - \sum_{t=1}^{T} p_{k_t} - TR(\frac{1}{T}\sum_{t=1}^{T} a_t x_t). \tag{3.1}$$

The penalty function $R(.)$ is a convex penalty function that measures the fairness cost of the decision-making process, due to the unbalancedness of the selected users at the end of the T rounds. It can be used to represent statistical parity [KAS11], as a measure of how far the allocation is from this fairness notion. This fairness penalty $R(.)$ is also used in [BLM21; Cel+22]. In fact, the objective (3.1) is equal the one used in these papers minus the cost of additional information bought, $\sum_t p_{k_t}$.

Knowns, unknowns and stochasticity We assume that users are i.i.d., in the sense that the whole vectors $(z_t, u_t, a_t, c_{t1} \ldots c_{tK})$ are i.i.d., drawn from some underlying unknown probability distribution. While this may be a strong assumption, some applications such as online advertising correspond to large T but to a short real time-frame, hence incurring very little variation in the underlying distribution. The prices p_k and the penalty function $R(.)$ are known beforehand. As mentioned several times, the only feedback received is c_{tk}, after selecting source k, and this should be enough to estimate u_t and a_t. We therefore assume that the conditional expectations $\mathbb{E}[u_t \mid c_{tk_t}]$ and $\mathbb{E}[a_t \mid c_{tk_t}]$ are known. The rationale behind this assumption is that these conditional expectations have been learned from past data.

Penalty examples and generalizations A typical example for a_t is the case of one-hot encoding: there are d protected categories of users and a_t indicates the category of user t. For simplicity, assume that $d = 2$, then the quantity $\sum_t a_{ti} x_t$ is the number of users of type $i \in \{1, 2\}$ that have been selected. The choice of $TR(\sum_{t=1}^{T} a_t x_t / T) = |\sum_t a_{t1} x_t - \sum_t a_{t2} x_t|$ amounts to penalizing the decision maker proportional to the absolute difference of users in both groups. This generic setting can also model other notions of fairness, such as Equality of Opportunity [HPS16], by choosing other values for a_t and $R(.)$, see examples and discussion in section 3.5.

Similarly, the choice of $x_t \in \{0, 1\}$ can be immediately generalized to any finite decision set or even continuous compact one (say, the reward at stage t would then be $\mathbf{u}_t^\top x_t$ for some vector $\mathbf{u}_t$), which makes it possible to handle problems such as bipartite matching. Instead of deriving a linear utility from selecting an user, general bounded upper semi-continuous (u.s.c.) utility functions can also be treated, and can be used to instantiate auctions mechanism (with some limitations detailed in section 3.7). We also explain how to relax the assumption that $\mathbb{E}[u_t \mid c_{tk_t}]$ is known in section 3.4.2, by deriving an algorithm that actually learns it in an online fashion, following linear contextual bandit techniques [APS11].

We show in section 3.3.5 that the assumption that $\mathcal{Z}$ is finite can be relaxed if all conditional distributions depend smoothly on z, following techniques from [PR13].

Mathematical assumptions We shall assume $|u_t| \leqslant \bar{u}$, for all t, and that $\|a\|_2 \leq 1$ for all $a \in \mathcal{A}$. We make, for now, no structural assumption on the variables c_{tk}. We mention here that the decision maker has to choose a single source at each stage, but this is obviously without loss of generality (by adding void or combination of sources).

We define $\Delta = \mathrm{Conv}(\mathcal{A} \cup \{0\})$, where Conv is the closed convex hull of a set. Since $\mathcal{A}$ is compact, the set Δ is also convex and compact. The penalty function $R : \Delta \to \mathbb{R}$ is a proper closed convex function that is L-Lipschitz continuous for the Euclidean norm $\|\cdot\|_2$. While the convexity assumption is pretty usual, the Lipschitzness assumption is rather mild as Δ is convex and compact. Nevertheless, interesting non-Lipschitz functions, such as the Kullback-Leibler divergence, can be modified to respect these assumptions (see [Cel+22]).

3.2.2 Benchmark and regret

A usual measure of performance for online algorithms is the regret that compares the utility obtained by an online allocation to the one obtained by an oracle that knows all parameters of the problem, yet not the realized sequence of private attributes. We denote the performance of this oracle by OPT:

$$\mathrm{OPT} = \max_{\boldsymbol{h} \in ([K]^{\mathcal{Z}})^T} \mathbb{E}\left[\max_{\boldsymbol{x} \in \{0,1\}^T} \mathbb{E}\left[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid c_{1k_1}, \ldots, c_{Tk_T}\right]\right], \tag{3.2}$$

where the conditional expectation indicates that the oracle first chooses a contextual policy h_t for all users that specifies which source $k_t = h_t(z_t)$ to select as a function of the variables z_t. It then observes all contexts c_{tk_t} and makes for all t the decisions x_t based on that.

Denoting ALG the expected penalized utility of an online algorithm, its expected regret is:

$$\mathrm{Reg} = \mathrm{OPT} - \mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x})] = \mathrm{OPT} - \mathrm{ALG}\,. \tag{3.3}$$

We remark that the benchmark of Equation (3.2) allows choosing different sources of information for different users with the same public information z_t. As such, it differs from classical benchmarks in the contextual bandit literature that compare the performance of an algorithm to the best static choice of arm per context and whose performance would be

$$\text{static}-\text{OPT} := \max_{h \in [K]^{\mathcal{Z}}} \mathbb{E}[\max_{\boldsymbol{x} \in \{0,1\}^T} \mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) | c_{1k_1}, \ldots, c_{Tk_T}]], \tag{3.4}$$

where the $h \in [K]^{\mathcal{Z}}$ policy that maps the public information to a source selection is the same for all users.

The benchmark (3.4) is the typical benchmark in contextual multi-armed bandits, as the global impact of decisions at different epochs and for different contexts are independent. However, this is no longer the case with the unfairness penalty $R(.)$ that requires coupling all decisions:

Proposition 3.1. *There exist an instance of the problem and a constant $b > 0$ such that for all T:*

$$\text{static}-\text{OPT} + bT < \text{OPT}\,.$$

This result shows that an algorithm that only tries to identify the best source will have a linear regret compared to OPT. This indicates that the problem of source selection and fairness are strongly coupled, even without public information available, and cannot be solved through some sort of two-phase algorithm where each problem is solved separately. The proof of this result is presented in section 3.10.1. In this paper, our primary emphasis lies in the examination of the performance disparity between an online algorithm and the offline optimum. Nevertheless, we provide supplementary experiments in section 3.10.3 that investigate how variations in the prices p_k and the penalty R impact the solution of the offline optimum.

3.3 Algorithm and Regret Bounds

In this part, we present our online allocation algorithm and its regret bound. For clarity of exposition, we first present the algorithm in the case $|\mathcal{Z}| = 1$, i.e., without public information available (and thus we remove z_t from the algorithm). The extension to $|\mathcal{Z}| > 1$ uses similar arguments and is discussed in section 3.3.5.

3.3.1 Overview of the algorithm

algorithm 1 devised to solve this problem is composed of two parts: a bandit algorithm for the source selection, and a gradient descent to adjust the penalty regularization term. This requires a dual parameter $\lambda_t \in \mathbb{R}^d$ that is used to perform the source selection as well as the allocation decision x_t.

The intuition is the following: the performance of each source is evaluated through some "dual value" for a given dual parameter λ. The optimal primal performance is equal to the dual value when it is minimized in λ, because R is convex and randomized combinations of sources is allowed thus there is no duality gap. Hence the dual value of each source is iteratively evaluated in order to select the best source, and simultaneously minimize the dual value of the selected source through λ, so that the source selected is indeed optimal.

Bandit part. For the source selection, we use the EXP3 algorithm (see Chapter 11 of [LS20a]) on a virtual reward that depends on the dual parameter. Given a dual parameter $\lambda \in \mathbb{R}^d$ and a context c_{tk}, we define the virtual reward as

$$\varphi(\lambda_t, c_{tk}, k) = \max\left(\mathbb{E}[u_t \mid c_{tk}] - \langle \lambda_t, \mathbb{E}[a_t \mid c_{tk}] \rangle, 0\right) - p_k, \tag{3.5}$$

where $\langle \lambda_t, \mathbb{E}[a_t \mid c_{tk}] \rangle$ denotes the scalar product between λ_t and $\mathbb{E}[a_t \mid c_{tk}]$. To compute this expectation, one needs to know the quantities p_k, $\mathbb{E}[u_t|c_{tk}]$ and $\mathbb{E}[a_t|c_{tk}]$.

To apply EXP3, a key property is to ensure that the virtual rewards are bounded, which requires λ_t to remain bounded. As we show in lemma 3.3, this is actually guaranteed by the design of the gradient descent on λ. This lemma implies that there exists $m \in \mathbb{R}^+$ such that $|\varphi(\lambda_t, c_{tk})| \leq m$ for all t and k. Let us denote by π_{tk} the probability that source k is chosen at time t and $k_t \sim \pi_t$. We define the importance-weighted unbiased estimator vector $\widehat{\varphi}(\lambda, c_t) \in \mathbb{R}^K$ where each coordinate $k' \in [K]$ is:

$$\widehat{\varphi}(\lambda, c_{tk}, k, k') = m - \mathbb{1}[k' = k]\frac{m - \varphi(\lambda, c_{tk}, k)}{\pi_{tk}}.$$

Using this unbiased estimator, we can apply the EXP3 algorithm to this virtual reward function.

Gradient descent part. Once the source k_t for user t is chosen and c_{tk_t} is observed, we can compute the decision x_t (see (3.6)). This x_t is then used in (3.8)-(3.9) to perform a dual descent step on the multiplier λ_t. Although using a dual descent step is classical, our implementation is different because we need to guarantee that the values of λ_t remain bounded for EXP3. To do so, we modify the geometry of the convex optimization sub-problem by considering a set of allocation targets larger than the original Δ. For $\delta \in \Delta$, we define the set Δ_δ as the ball of center δ and radius $\mathrm{Diam}(\Delta)$. This ball contains Δ: $\Delta \subset \Delta_\delta$.

algorithm 1 uses any extension of R to $\bar{\Delta} = \cup_{\delta \in \Delta} \Delta_\delta$ that is convex and Lipschitz-continuous, for instance, the following one (see lemma 3.10):

$$\bar{R}(\delta) = \inf_{\delta' \in \Delta} \{R(\delta') + L\|\delta - \delta'\|_2\},$$

that has the same Lipschitz-constant L as R (which is the best Lipschitz-constant possible).

3.3.2 Algorithm and implementation

Combining these different ideas leads to algorithm 1. This algorithm maintains a dual parameter λ_t that encodes the history of unfairness that ensures that the fairness penalty $R(.)$ is taken into account. This dual parameter λ_t is used in (3.6) to compute the allocation x_t and in (3.7) to choose the source of information. The dual update (3.8)-(3.9) guarantees that we take $R(.)$ into account.

An interesting property of the dual gradient descent is that it manages to provide a good fair allocation while updating the source selection parameters simultaneously. Indeed if λ were fixed, then we could solve the source selection part through φ and a bandit algorithm. However, both the λ_t and the π_t change over time, which may hint at the necessity of a two-phased algorithm as the combination of these two problems generates non-stationarity for both the dual update and also for the bandit problem. The dual gradient descent manages to combine both updates in a single-phased algorithm.

The different assumptions imply that the decision maker has access to the $\mathbb{E}[a_t|c_{tk}]$ and to $\mathbb{E}[u_t \mid c_{tk}]$ for any possible context value c_{tk}. Such values could be indeed estimated from offline data. The knowledge of such values is sufficient to compute the allocation x_t in (3.6), the virtual value estimation $\widehat{\varphi}$ of (3.7), or to compute δ_t. Once these values are computed, the only difficulty is to solve (3.8). In some cases, it might be solved analytically. Otherwise, it can also be solved numerically as it is an (a priori low-dimensional) convex optimization problem. Overall this is an efficient online algorithm which only uses the current algorithm parameters λ_t, π_t, and current context c_{tk_t}.

Algorithm 1 Online Fair Allocation with Source Selection

Input: Initial dual parameter λ_0, initial source-selection-distribution $\pi_0 = (1/K, \ldots, 1/K)$, step sizes η and ρ, cumulative estimated rewards $S_0 = 0 \in \mathbb{R}^K$.

for $t \in [T]$ **do**

Draw a source $k_t \sim \pi_t$ where $\pi_{tk} = \exp(\rho S_{t-1,k}) / \sum_{l=1}^{K} \exp(\rho S_{t-1,l})$, and observe c_{tk_t}.

Compute the allocation for user t:

$$x_t = \begin{cases} 1 & \text{if } \mathbb{E}[u_t \mid c_{tk_t}] \geqslant \langle \lambda_t, \mathbb{E}[a_t \mid c_{tk_t}] \rangle \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

Update the estimated rewards sum and sources distributions for all $k \in [K]$:

$$S_{tk} = S_{(t-1)k} + \widehat{\varphi}(\lambda_t, c_{tk}, k_t, k), \tag{3.7}$$

Let $\delta_t = x_t \mathbb{E}[a_t \mid c_{tk_t}]$. Compute the dual fairness allocation target and update the dual parameter

$$\gamma_t = \underset{\gamma \in \Delta_{\delta_t}}{\arg\max} \{ \langle \lambda_t, \gamma \rangle - \bar{R}(\gamma) \}, \tag{3.8}$$

$$\lambda_{t+1} = \lambda_t - \eta(\gamma_t - \delta_t). \tag{3.9}$$

end for

3.3.3 Regret bound

We emphasize that algorithm 1 uses randomization among the different sources of information and does not aim to identify the best source. As shown in proposition 3.1, this is important because using the best source of information can be strictly less good that using randomization. Moreover, it simplifies the analysis because it convexifies the set of strategies. This means that Sion's minimax theorem can be applied to some dual function, which allows for λ_t and π_t to be updated simultaneously. If one would try to target the best static source of information (static-OPT), one would need to determine the optimal dual parameter λ_k of each source. This would lead to an algorithm that is both more complicated and less efficient (because of proposition 3.1).

The following theorem shows that algorithm 1 has a sub-linear regret of order $O(\sqrt{T})$. This regret bound is comparable to those in [BLM21; Cel+22] but we handle the much more challenging case of having multiple sources of information, and imperfect information about a_t and u_t.

Theorem 3.2. *Assume that Algorithm 1 is run with the parameters* $\eta = L/(2\,\mathrm{Diam}(\Delta)\sqrt{T})$, $m = \bar{u} + L + \max_k |p_k| + 2\eta\,\mathrm{Diam}(\Delta)$, $\rho = \sqrt{\log(K)/(TKm^2)}$, *and that* $\lambda_0 \in \partial R(0)$, *the subgradient of R at 0. Then the expected regret of the algorithm is upper bounded by:*

$$\mathrm{Reg} \leqslant 2((L + \bar{u} + \max_k |p_k|)\sqrt{K \log(K)} + L\sqrt{d} + L\,\mathrm{Diam}(\Delta))\sqrt{T} + 2L\sqrt{K \log(K)}.$$

Note that this regret bound is tight: when $R = 0$ the problem we consider reduces to a K-armed bandit with bounded rewards $\mathbb{E}[\max(\mathbb{E}[u_t \mid c_{t,k}], 0) - p_k]$ for arm $k \in [K]$, which has a $\Omega(\sqrt{T})$ regret lower bound [LS20a]. It is of the same order in T as our regret upper bound. Remark that $R = 0$ implies that $L = 0$ and we do recover the regret bound for the EXP3 algorithm. Similarly if $K = 1$ the bandit regret contribution disappears. In our analysis, the regret due to the interaction between the bandit and the online fair allocation is $L\sqrt{K \log(K) T}$.

The time-horizon dependent parameters used in algorithm 1 can be adapted to obtain an anytime algorithm. While using a doubling trick directly for η is not possible as some protected attribute would already be selected when restarting the algorithm, we can use an adaptive learning rate of $\eta_t = \mathcal{O}(1/\sqrt{t})$. Indeed due to the boundedness of the $(\lambda_t)_{t \in T}$ (lemma 3.3), we can act as if we had a finite diameter for the space of the λ_t. This results in a slight increase of regret, the constant term in theorem 3.2 now scaling in $\mathcal{O}(\sqrt{T})$.

3.3.4 Sketch of proof

As mentioned above, when $K = 1$ (resp. $R = 0$) the problem reduces to fair online allocation as in [BLM21; Cel+22] (resp. to multi-armed bandits). The main technical difficulty thus lie in combining algorithms used in these problems. Ideally, we would have access to some optimal dual parameter λ_k^* *before* we run the algorithm so that we can simply run a bandit algorithm, which is obviously not possible as the selected source k_t affects the fairness, and the selected parameter λ_t affects the arms virtual rewards. In particular it is not clear how the λ_t evolve in the worst case while the algorithm is running. Instead we alternate between those primal and dual updates. We thus need to show that this alternation indeed achieves good performance.

We present two important lemmas used in the proof of theorem 3.2. The first one guarantees that doing a gradient descent with $\bar{R}$ implies that the dual values λ_t remain bounded. This is crucial for EXP3 as it implies that the virtual values φ remain bounded along the path of the λ_t.

Lemma 3.3. *Let $\lambda_0 \in \mathbb{R}^d$, $\eta > 0$, and an arbitrary sequences $(\delta_1, \delta_2, \dots) \in \Delta^{\mathbb{N}}$. Assume that $\lambda_0 \in \partial R(0)$ and define recursively λ_t by eqs.* (3.8) *and* (3.9)*. Then for all t, we have $\|\lambda_t\|_2 \leqslant L + 2\eta \operatorname{Diam}(\Delta)$.*

Sketch of Proof. The main idea is to show that the distance of λ_t to the reunion of subgradients of $\bar{R}$ (which is bounded by Lipschitzness property) is a decreasing function in t. We can show this by using the KKT conditions of the optimization problem. The convex set over which we optimize is a simple Euclidean ball, because of our modification, centered around the appropriate point hence allowing us to redirect the gradient $\lambda_{t+1} - \lambda_t$ towards this set. Moreover, we add some "security" around this set of the size of the gradient bound to make sure that the λ_t remains in this set. The full proof can be found in section 3.8.3. □

The second lemma guarantees that having access to the conditional expectation $\mathbb{E}[a_t|c_{tk}]$ is enough to derive a good algorithm when considering the conditional expectation of the total utility. This way we avoid the computation of the conditional expectation of R, which would be more difficult.

Lemma 3.4. *For $(x_1, \dots, x_T)$ and $(k_1, \dots, k_T)$ generated according to algorithm 1, with $\delta_t = x_t \mathbb{E}[a_t \mid c_{tk_t}]$, we have the following upper bound:*

$$\left|\mathbb{E}[R(\frac{1}{T}\sum_{t=1}^{T} a_t x_t) - R(\frac{1}{T}\sum_{t=1}^{T}\delta_t)]\right| \leqslant 2L\sqrt{\frac{d}{T}}.$$

Sketch of proof. We use the Lipschitz property of R and some inequalities to directly compare the difference of the sums. The variable x_t depends on the past history and needs to be carefully taken into account, through proper conditioning. Finally, we compute the variance of a sum of martingale differences. See proof in section 3.8.4. □

Using these two Lemmas, we now give the main ideas of the proof of theorem 3.2. First, we upper-bound OPT through a dual function involving the convex conjugate R^*, with similar arguments as in [BLM21; Cel+22]. Then we need to lower-bound ALG with this dual function minus the regret. The main difficulty is that the source virtual rewards distribution changes with λ_t, and so does the average λ_t target through π_t. The performance of ALG can be decomposed at each step t into the sum of the virtual reward and a term in $R^*(\lambda_t)$ encoding the fairness penalty. We deal with the virtual rewards using an adversarial bandits algorithm able to handle any reward sequence using techniques for adversarial bandits algorithm from [Haz22;

LS20a], as the λ_t are generated by a quite complicated Markov-Chain. These rewards are bounded because of lemma 3.3. This yields the two regret terms in $\sqrt{K\log(K)}$, the second one stems from the difference between $\bar{\Delta}$ and Δ. For the fairness penalty, using the online gradient descent on the λ_t we end up being close to the penalty of the conditional expectations up to the regret term in $\mathrm{Diam}(\Delta)$. This last term is close to the true penalty through lemma 3.4. This provides us with a computable regret bound where all the parameters are known beforehand. The full proof can be found in section 3.9.

3.3.5 Public contexts

We now go back to the general case with $|\mathcal{Z}| > 1$ finite. We would like to derive a good algorithm, which also takes into account the public information z_t. Reusing the analogy with bandit problems, this seems to be akin to the contextual bandit problem, where we would simply run the algorithm for each context in parallel. However, this would be incorrect: not only for a fixed public context does the non-separability of R couples the source selection with the fairness penalty, it also couples all of the public contexts z_t together. Hence the optimal policy is not to take the best policy for each context, which is once again different from the classical bandit setting.

The solution is to run an anytime version of the EXP3 algorithm for each public context in $\mathcal{Z}$, but to keep a common λ_t for the evaluation of the virtual value. While it is technically not much more difficult and uses similar ideas to what was done for different sources when $|\mathcal{Z}| = 1$, the fact that only one of the "block" of the algorithm (the bandit part) needs to be parallelized, is quite specific and surprisingly simple.

Proposition 3.5. *For μ the probability distribution over $\mathcal{Z}$ finite, we can derive an algorithm that has a modified regret of order $\mathcal{O}(\sqrt{TK\log(K)}\sum_{z\in\mathcal{Z}}\sqrt{\mu(z)})$, where $\mu(z)$ is the probability that the public attribute is z.*

The algorithm and its analysis are provided in section 3.14.1. Note that in the worst case, when $\mathcal{Z}$ is finite, one has $\sum_{z\in\mathcal{Z}}\sqrt{\mu(z)} \leqslant \sqrt{|\mathcal{Z}|}$ by Jensen's inequality on the square root function.

If $\mathcal{Z}$ is not finite but a bounded subset of a vector space set of dimension r, such as $\mathcal{Z} = [0,1]^r$, additional assumptions on the smoothness of the conditional expectations are sufficient to obtain a sub-linear regret algorithm:

Proposition 3.6. *If the conditional expectations of a_t and u_t are both Lipschitz in z, then discretizing the space $\mathcal{Z} = [0,1]^r$ through an ε-cover and applying the previous algorithm considering that one public context corresponds to one of the discretized bins, we can obtain a regret bound of order $\mathcal{O}(T^{(r+1)/(r+2)})$.*

The proof of this result uses standard discretization arguments under Lipschitz assumptions from [PR13], which can also be found in Chapter 8.2 of [Sli19] or in Exercise 19.5 of [LS20a]. Specificities for this problem, such as these assumptions being enough to guarantee that φ is Lipschitz in z, can be found in section 3.14.2.

3.4 Extensions

3.4.1 Other types of fairness penalty

The fairness penalty term of eq. (3.1) is quantified as $TR(\sum a_t x_t / T)$. While this term is the same as the one used in [Cel+22] and can encode various fairness definitions (see the discussion in the aforementioned paper), this does not encompass all possible fairness notions. For instance, one may want to express fairness as a function of $\sum a_t x_t / \sum_t x_t$, which is the conditional empirical distribution of the user's protected attributes given that they were selected. This would lead to replacing the original penalty term by $(\sum_t x_t) R(\sum a_t x_t / \sum_t x_t)$.

As pointed out in [Cel+22], one possible issue is that $R(\sum a_t x_t / \sum_t x_t)$, in general, is not convex in x, even if R is convex. However $(\sum_t x_t) R(\sum a_t x_t / \sum_t x_t)$ is the perspective function of $R(A(.))$ (with A the matrix with columns the a_t), which is thus convex. Hence, algorithm 1 can be adapted to handle this new fairness penalty with two modifications. First, for the bandit part, we run the algorithm with a new virtual reward function, expressed as:

$$\tilde{\varphi}(\lambda_t, c_{tk}, k) = \max(\mathbb{E}[u_t \mid c_{tk}] - \langle \lambda_t, \mathbb{E}[a_t \mid c_{tk}] \rangle + R^*(\lambda_t), 0) - p_k.$$

This leads to an allocation $x_t = 1$ if $\mathbb{E}[u_t \mid c_{tk_t}] + R^*(\lambda_t) \geqslant \langle \lambda_t, \mathbb{E}[a_t \mid c_{tk_t}] \rangle$ and $x_t = 0$ otherwise.

Second, we modify the set on which the dual descent is done. The set Δ now becomes $\tilde{\Delta} = \mathrm{Conv}(\mathcal{A})$ (without the union with 0), and we now use $\tilde{\delta}_t = \mathbb{E}[a_t \mid c_{tk_t}]$ instead of $\delta_t = x_t \mathbb{E}[a_t \mid c_{tk_t}]$. Line (3.8) remains unchanged up to replacing δ and Δ by $\tilde{\delta}$ and $\tilde{\Delta}$. Finally the dual parameter update now becomes $\lambda_{t+1} = \lambda_t - \eta x_t(\gamma_t - \tilde{\delta}_t)$, which means that whenever $x_t = 0$, λ_t does not change.

With these modifications, the following theorem (whose proof is very similar to the one of theorem 3.2 and detailed in section 3.11), shows that we recover similar regret bounds as previously, with some modified constants due to the presence of R^* in the virtual reward, and the modified Δ. This yields a new class of usable fairness penalties which was previously not known to work.

Theorem 3.7. *Using algorithm 1 with the modifications detailed above, the regret with respect to the objective with the modified penalty* $(\sum_t x_t)R(\sum_t a_t x_t / \sum_t x_t)$ *is of order* $\mathcal{O}(\sqrt{T})$.

3.4.2 Learning conditional utilities $\mathbb{E}[u_t \mid c_{tk}]$

To compute the virtual values, the algorithm relies on the knowledge of $\mathbb{E}[u_t \mid c_{tk}]$ for all possible context values. We now show how to relax this assumption even if the decision maker receives the feedback u_t only when the user is selected ($x_t = 1$). We shall make the usual structural assumption of the classical stochastic linear bandit model. The analysis relies on Chapters 19 and 20 of [LS20a], and only the main ideas are given here. For simplicity we will assume that there are no public contexts ($|\mathcal{Z}| = 1$).

We assume that the contexts c_{tk} are now feature vectors of dimension q_k, and that for all $k \in [K]$, there exists some vector $\psi_k \in \mathbb{R}^{q_k}$ so that

$$u_t - \langle \psi_k, c_{tk} \rangle, \tag{3.10}$$

is a zero mean 1-subgaussian random variable conditioned on the previous observation (see a more precise definition in section 3.12). We make classical boundedness assumptions, that for all k: $\|\psi_k\|_2 \leqslant \bar{\psi}$, that $\|c_{tk}\|_2 \leqslant \bar{c}$, and that $\langle \psi_k, c_{tk} \rangle \leqslant 1$ for all possible values of c_{tk}.

Intuitively, we aim at running a stochastic linear bandit algorithm on each of the different sources for T_k steps, where T_k is the number of times that source k is selected, but this breaks because of dependencies among the k sources. Hence, what we do is to slightly change the rewards and contextual actions, so that we do something akin to artificially running K bandits in parallel for T steps. We force a reward and action of 0 for source k whenever it is not selected, and consider that each source is run for T steps (even if it is actually selected less than T times). Thus we can directly leverage and modify the existing analysis for the stochastic linear bandit. The full algorithm is given in section 3.12.

Proposition 3.8. *Given the above assumptions, the* added *regret of having to learn* $\mathbb{E}[u_t \mid c_{tk}]$ *is of order* $\mathcal{O}(\sqrt{KT}\log(T))$.

If the decision maker knows the optimal combination of sources, she can simply select sources proportionally to this combination, and actually learn the conditional expectation independently for each source. The worst case is to have to pull each arm T/K times, which then incurs an additional regret with the same $\sqrt{K}$ constant. This shows that this bound is actually not too wasteful, as the cost of artificially running these bandits algorithm in parallel only impacts the logarithmic term.

Conclusion

We have shown how the problem of optimal data purchase for fair allocations can be tackled, using techniques from online convex optimization and bandits algorithms. We proposed a computationally efficient algorithm yielding a $\mathcal{O}(\sqrt{T})$ regret algorithm in the most general case. Interestingly, because of the non separable penalty R, the benchmark is different from a bandit algorithm, as randomization can strictly improve the performance for some instances, even though the setting is stochastic. We have also presented different types of fairness penalties that we can additionally tackle compared to previous works, in particular in the full information setting, and some instances where assumptions on the decision maker's knowledge can be relaxed.

Throughout the paper (even in section 3.4.2 where we relax the assumption that $\mathbb{E}[u_t \mid c_{tk}]$ is known), we assumed that the decision-maker knows $\mathbb{E}[a_t \mid c_{tk}]$. This assumption is reasonable as this is related to demographic data and not to a utility that may be specific to the decision maker. Yet, this assumption can also be relaxed in specific cases. As an example, suppose that the decision maker can pay different prices to receive more or less noisy versions of a_t that correspond to the data sources (the pricing could be related to different levels of Local Differential Privacy for the users, see [DR14]). Then, under some assumptions, $\mathbb{E}[a_t \mid c_{tk}]$ can also be learned online—we defer the details to section 3.13.

The works of [BLM21] and [Cel+22] can include hard constraints on some budget consumption when making the allocation decision x_t, which we did not include in our work. If this budget cost is measurable with respect to c_{tk}, then our analysis does not preclude using similar stopping time arguments as was done in these two works. However if this budget consumption is completely random, our algorithm and analysis can not be directly applied as this budget consumption may give additional information on the a_t that was just observed. Regarding the i.i.d. assumption, in [BLM21] adversarial inputs are also considered, and the same could be done here for the contexts $c_{t,k}$ with similar results. However some stochasticity is still needed so that $\mathbb{E}[u_t \mid c_{t,k}]$ is well defined, which leads to an ambiguous stochastic-adversarial model. An open question would be to consider an intermediate case where $\mathbb{E}[u_t \mid c_{t,k}]$ would depend on t, and take into account learning for non-stationary distributions.

3.5 Fairness Penalty

In the main part of the paper, we have considered a penalty with a general form $R(\sum_t a_t x_t / T)$. Here we give examples of fairness notions that can be encoded through this general penalty.

Statistical parity Notions likes statistical parity are defined as constraints on the independence between protected attributes and selection decisions. More precisely, if X denotes a treatment variable, and A the protected attributes of an individual, we say that X and A are fair with respect to statistical parity if X and A are independent.

If $X \in \mathcal{X}$ and $A \in \mathcal{A}$ can take a finite number of values, this can be equivalently rewritten using probability distribution: X satisfies statistical parity if for all $(x, a) \in \mathcal{X} \times \mathcal{A}$:

$$\mathbb{P}(A = a, X = x) = \mathbb{P}(A = a)\mathbb{P}(X = x). \tag{3.11}$$

To encode statistical parity in our framework, we consider that the a_t are one-hot encoding of the protected attributes (*i.e*, a_t is a vector with $d = |\mathcal{A}|$ dimensions that is equal to 0 everywhere except on the coordinate corresponding to the protected attribute of individual t, that equals 1). For the decision variables $x_t \in \{0, 1\}$, $\sum_t x_t$ is the number of individual selected and $\sum_t (a_t)_i x_t$ is the number of selected individual whose protected attributed is $i \in [d]$. Hence, (3.11) rewrites as: for all $i \in [d]$:

$$\frac{\sum_t (a_t)_i x_t}{T} = \left(\frac{\sum_t (a_t)_i}{T}\right)\left(\frac{\sum_t x_t}{T}\right).$$

Recall that the variables a_t are *i.i.d.* and let us denote by $\alpha \in \mathcal{P}_d$ the probability distribution vector of the a_t. For large T, we have $\sum_t a_t / T \approx \alpha$. Which means that we will say that an allocation vector $\boldsymbol{x}$ satisfies statistical parity if $\sum_t a_t x_t/T = \sum_t \alpha x_t/T$, which can be equivalently rewritten as

$$\frac{\sum_t a'_t x_t}{T} = 0,$$

with $a'_t = a_t - \alpha$.

Penalties R corresponding to statistical parity In practice, achieving strict statistical parity is very constraining, and most of the time some relaxation of this constraint is allowed, at the price of some penalty [Aha+21]. In our paper, we penalize the non-independence through the function R. There are different choices of penalty function that can be used:

- A natural choice can be to penalize as a function of the distance between joint empirical distribution $a'_t x_t/T$ and 0. For instance, one could use $R(a'_t x_t/T) = ||a'_t x_t/T||^\beta$ for any given norm $||\cdot||$ and any parameter $\beta > 0$.
- It is also possible to measure the non-independence with the empirical distribution of the x_t conditioned on $a_t = a$, which should be equal regardless of the protected group $a \in \mathcal{A}$. In this case, the empirical distribution conditioned on the protected group $i \in [d]$ is $\sum_t a_{ti} x_t / \sum_t a_{ti} \approx \sum_t a_{ti} x_t/(\alpha_i T)$. Similarly we the previous case, R can penalize how far is this vector from 0.

In the two cases, the fairness penalty can be written as $TR(\sum_t a'_t x_t/T)$, for some modified a'_t that are not necessarily one-hot encoder. This explains why we choose the set $\mathcal{A}$ to be more general than the set of one-hot encoders.

If we want to target another specific proportion for each protected group, we can replace α by any other distribution ν. Note that another possibility is to instead use the empirical distribution of the a_t conditioned on the x_t. This time the form of the vector is slightly different, it is $\sum_t a_t x_t / \sum_t x_t$; this is the type of penalty tackled and discussed in section 3.4.1.

Equality of opportunity Another popular fairness notion is the one of equality of opportunity [HPS16], where the probability distribution is further conditioned on another finite random variable y_t taking some discrete value in $\mathcal{Y}$ that indicates some fitness of the individual. For Y a random variable with an outcome $y \in \mathcal{Y}$

considered as a positive opportunity, we say that X, A and Y are fair with respect to equality of opportunity, if

$$X \text{ and } A \text{ are independent given } [Y = y].$$

In order to adapt this notion in the online allocation setting, is sufficient to consider the cartesian product $\mathcal{A}' = \mathcal{A} \times \mathcal{Y}$ as an extended new set. Then if we would like to penalize with a metric related to equality of opportunity, we can for instance compose R with the projection over the coordinates corresponding to $y_t = 1$.

Overall, this shows that the choice of R, the type of fairness vectors, and the choice of $\mathcal{A}$ are all crucial when designing a fairness penalty.

3.6 A More General Online Allocation Setting

In the main body of the paper, we choose to simplify the exposition and to restrict ourselves to $x_t \in \{0, 1\}$ and to formulate utilities as scalar variables u_t. In the following, the regret bounds of algorithm 1 are obtained in a setting that is more general than the one previously introduced:

- We now consider that the allocations decisions x_t are in some set $\mathcal{X} \subset \mathbb{R}^n_+$. This makes it possible to tackle bipartite matching problems, by letting $\mathcal{X} = \{x \in \mathbb{R}^n_+ \mid \sum_{i=1}^n x_i \leqslant 1\}$ be a matching polytope.
- Instead of a random scalar u_t being drawn, we suppose that a random function $f_t : \mathcal{X} \to \mathbb{R}$ is drawn. The functions f_t are supposed u.s.c. and are all bounded in absolute value by some $\bar{f}$. This could simply be a deterministic concave function of u_t for instance. Remark that when we assume that the conditional expectation can be computed, the decision maker's required knowledge is more demanding, as she needs to be able to compute it for all $x \in \mathcal{X}$.
- Instead of a random vector, a continuous random function $a_t : \mathcal{X} \to \mathcal{A}$ can be drawn, with $\|a_t(x)\|_2$ uniformly bounded for all a_t and x, and we define $\Delta = \mathrm{Conv}(\mathcal{A})$. We suppose that $\|a_t(x)\|_2 \leqslant 1$, which can be done without loss of generality.

The setting of section 3.2.1 is recovered by using $f_t(x) = u_t x$, $a_t(x) = a_t x$ and $\mathcal{X} = \{0, 1\}$. For the more general setting, we need to redefine the virtual value with f_t and a_t as:

$$\varphi(\lambda, c_{tk}, k) = \max_{x \in \mathcal{X}} \mathbb{E}[f_t(x) - \langle \lambda_t, a_t(x) \rangle \mid c_{tk}] - p_k \tag{3.12}$$

This virtual reward is well defined by lemma 3.9. We rewrite algorithm 1 using f_t and this new virtual value to obtain line 0. There are two differences with algorithm 1: First we use the argmax function in (3.13) instead of the explicit form of eq. (3.6); Second we use the virtual value function (3.12) in eq. (3.14) instead of the virtual value function of eq. (3.5).

Algorithm 2 Online Fair Allocation with Source Selection — General algorithm

Input: Initial dual parameter λ_0, initial source-selection-distribution $\pi_0 = (1/K, \ldots, 1/K)$, dual gradient descent step size η, EXP3 step size ρ, cumulative estimated rewards $S_0 = 0 \in \mathbb{R}^K$.

for $t \in [T]$ **do**

Draw a source $k_t \sim \pi_t$, where $(\pi_t)_k \propto \exp(\rho S_{(t-1)k})$ and observe c_{tk_t}.

Compute the allocation for user t:

$$x_t = \arg\max_{x \in \mathcal{X}} \mathbb{E}[f_t(x) - \langle \lambda_t, a_t(x) \rangle \mid c_{tk_t}]. \tag{3.13}$$

Update the estimated rewards sum and sources distributions for all $k \in [K]$:

$$S_{tk} = S_{t-1,k} + \widehat{\varphi}(\lambda_t, c_{tk}, k_t), \tag{3.14}$$

Compute the expected protected group allocation $\delta_t = \mathbb{E}[a_t(x_t) \mid c_{tk_t}, x_t]$ and compute the dual protected groups allocation target and update the dual parameter:

$$\gamma_t = \arg\max_{\gamma \in \Delta_{\delta_t}} \{\langle \lambda_t, \gamma \rangle - \bar{R}(\gamma)\},$$
$$\lambda_{t+1} = \lambda_t - \eta(\gamma_t - \delta_t).$$

end for

3.7 Reduction from auctions

In the rest of the paper, we assume that the decision variables are the values x_t. In this section, we explain how our setting can be adapted the the case of targeted advertisement, that are usually allocated through auction mechanisms. In the later, the decision variables are the bids but not the x_t: the variables x_t are a consequence of the bids of the decision maker and of the bids of others.

Model The model that we consider is as follows: for user t, after choosing a source of information k_t, the decision maker observes c_{tk_t} and decides to bid $b_t \in \mathcal{B}$. The allocation $x_t = 1$ is given if her bid is higher than the maximum bid of the other bidders. If the bidder wins the auction, it then pays a price that depends on the auction format (*e.g.*, b_t for first price auctions and the second highest bid for second price auctions). Assuming that the bid of the others are not affected by the decision maker's strategy, the bids of the others are *i.i.d.* (see [IJS11; BBW14; FJ17] for a discussion about the stationarity assumption). In this case, one falls back to our model with *i.i.d.* users, except that the decision variables are the variables b_t and we have $x_t(b_t)$. We can apply our algorithm to this setting, by replacing the set of strategies $\mathcal{X}$ by the set of possible bids $\mathcal{B}$, which is also compact. For that, we need to consider a new utility function $b \mapsto u_t(b)x_t(b)$ which is a bounded u.s.c. function, and the new protected attributes function $b \mapsto a_t x_t(b)$. If $\mathbb{E}[a_t x_t(b) \mid c_{tk}]$ is continuous in b, this falls under the setting that we described in section 3.6.

Examples of utilities: first-price and second price auctions Let us denote by v_t the value of the user t for the decision maker, and by d_t the highest bid of the others for this user. The decision maker will win the auction if $b_t \geqslant d_t$. Hence, the allocations will be $x_t(b_t) = \mathbb{1}_{b_t > d_t}$. Moreover, the utility function will be:

- $f_t(b_t) = (v_t - b_t)x_t(b_t)$ for first price auctions.
- $f_t(b_t) = (v_t - d_t)x_t(b_t)$ for second price auctions.

These functions are u.s.c with respect to b_t. Moreover, under mild assumption that the distribution of the highest bid of the others d_t has a density, the function $\mathbb{E}[a_t x_t(b) | c_{tk_t}]$ is continuous. Hence, our algorithm can be applied.

Difficulty: Knowledge is power To apply the algorithm, the main difficulty is to be able to compute the value b_t that maximizes eq. (3.12). For that it is sufficient for the decision maker to have access to $\mathbb{E}[f_t(b_t)|c_{tk_t}]$ and $\mathbb{E}[a_t x_t(b_t)|c_{t,k_t}]$. This presupposes a lot of distributional knowledge from her part. In the case of full information with a_t and v_t directly given, [BLM21; Cel+22] were able to obtain an efficient reduction from online allocations to second price auctions, using the incentive compatible nature of this kind of auction. Similarly, [FJ17] were able to derive efficient bidding strategies using limited knowledge for repeated auctions with budget constraints under some independence assumptions. Translating what is done in these two lines of work to our general model is difficult because we would need to rely on the assumption that a_t and d_t are independent conditioned on c_{tk_t}. While we think that this assumption is not reasonable because different agents might have access to different information sources, and because as the a_t is common to all bidders it is correlated with all the bids hence also with d_t, we explain in the remark below how assuming independence can lead to a simple algorithm for second price auctions.

Some remarks:

- For our model, the first price and second price auctions are equally difficult: The amount of knowledge needed to compute b_t and π_t is the same for both cases and first price auction is not more difficult than the second-price auction setting (contrary to what usually happens).

- During this whole reduction from auctions, we have not assumed that the d_t (or x_t) are observed as a feedback (whether the bidder wins the auction or not), and thus they are not assumed to be observed in the benchmark either. This is because observing d_t or x_t could give us access to information on a_t that are not contained in c_{tk_t}.

- If we do make the assumption that (u_t, a_t) is independent of d_t conditionally on all c_{tk}, then we do recover an easy algorithm for second price auctions, and we can allow for a benchmark taking into account the feedback d_t (simply because now this feedback does not give any additional information for the estimation of a_t). Indeed, in that case the optimal bid is $b_t = \mathbb{E}[v_t - \langle \lambda_t, a_t \rangle \mid c_{tk}]$. This follows from similar argument made in [FJ17], as for b_t some $\mathcal{H}_t$ measurable bidding strategy we have:

$$\begin{aligned}\mathbb{E}[v_t x_t - \langle \lambda_t, a_t \rangle x_t - d_t x_t \mid c_{tk_t}] = & \mathbb{E}[\mathbb{E}[v_t x_t - \langle \lambda_t, a_t \rangle x_t - d_t x_t \mid \mathcal{H}_t, d_t] \mid c_{tk_t}] \\ =& \mathbb{E}[\mathbb{E}[v_t \mid c_{tk_t}] x_t - \langle \lambda_t, \mathbb{E}[a_t \mid c_{tk_t}] \rangle x_t - d_t x_t \mid c_{tk}] \\ =& \mathbb{E}[x_t(\mathbb{E}[v_t \mid c_{tk_t}] - \langle \lambda_t, \mathbb{E}[a_t \mid c_{tk_t}] \rangle - d_t) \mid c_{tk}] \\ \leqslant & \mathbb{E}[\max(\mathbb{E}[v_t \mid c_{tk_t}] - \langle \lambda_t, \mathbb{E}[a_t \mid c_{tk_t}] \rangle - d_t, 0) \mid c_{tk}],\end{aligned}$$

  and this last inequality is reached for $b_t = \mathbb{E}[v_t - \langle \lambda_t, a_t \rangle \mid c_{tk}]$.

3.8 Technical & Useful Lemmas

In this section, we prove several lemmas that are useful for the proof of the main theorem 3.2.

3.8.1 Preliminary lemmas

We first provide some simple lemmas. For completeness, we provide proofs of all results, even if they are almost direct consequences of the definitions.

Lemma 3.9. *The virtual value function φ is well defined by Equation* (3.12)*, as the* $\arg\max$ *is non-empty and the maximum is reached.*

Proof. We simply need to have that $\mathbb{E}[f_t(x) \mid c_{tk}]$ and $-\langle \lambda, \mathbb{E}[a_t(x) \mid c_{tk}]\rangle$ are u.s.c. in x. The continuity of a_t is used to guarantee that $-\langle \lambda, a_t(x)\rangle$ is u.s.c. regardless of the values of λ. This is immediate from the definition of upper semi-continuity and the Reverse Fatou Lemma. We detail it here for f_t for completeness. We say that f_t is u.s.c. if and only if

$\forall x \in \mathcal{X}$ if there is a sequence $(x_n)_{n\in\mathbb{N}}$ so that $x_n \xrightarrow[n\to\infty]{} x$, then $\limsup_{n\to\infty} f_t(x_n) \leqslant f_t(x)$.

For $x \in \mathcal{X}$, let $(x_n)_{n\in\mathcal{N}}$ be a sequence so that $x_n \to x$. Let P be a probability distribution over f_t conditionally on c_{tk}. The function $|f_t(x_n)| \leqslant \bar{f}$, $f_t(x_n)$ is integrable, we can apply the Reverse Fatou Lemma for conditional expectation. If $f_t(x_n)$ is not positive, we can apply it to $f_t(x_n) + \bar{f}$ which is positive. Using the Reverse Fatou Lemma and the upper semi-continuity of f_t we obtain:

$$\limsup_{n\to\infty} \int f_t(x_n)dP \leqslant \int \limsup_{n\to\infty} f_t(x_n)dP \leqslant \int f_t(x)dP.$$

This means that $x \mapsto \mathbb{E}[f_t(x) \mid c_{tk}]$ is u.s.c. over $\mathcal{X}$.

Because $\mathcal{X}$ is compact, and because the maximum of an u.s.c. function is reached over a compact set, the maximum is well defined. □

Lemma 3.10. *Let* $\bar{R} : \delta \mapsto \inf_{\delta'\in\Delta}\{R(\delta') + L\|\delta - \delta'\|_2\}$*. Then* $\bar{R} = R$ *over* Δ*,* $Lip(R) = Lip(\bar{R})$*, and* $\bar{R}$ *is convex.*

Proof. The function $\bar{R}$ is obtained through the explicit formula of the Kirszbraun theorem, which states that we can extend functions over Hilbert spaces with the same Lipschitz constant. Hence the equality over Δ and the same Lipschitz constant.

It remains to prove the convexity. Note that because R is proper, it is lower semi-continuous, $\|\delta - \cdot\|_2$ as well, and Δ is compact, hence the $\inf$ is reached and it is actually a minimum. Let δ_1, δ_2 in $\mathbb{R}^d$ and $\alpha \in [0,1]$, and let u_1 and u_2 be the minimums of the optimization problem for respectively δ_1 and δ_2. Let $u =$

$\alpha u_1 + (1-\alpha)u_2$. We have $u \in \Delta$ because Δ is convex, and using that R is convex and we have

$$R(u)+L\|(\alpha\delta_1+(1-\alpha)\delta_2)-u\|_2 \leqslant \alpha(R(u_1)+L\|\delta_1-u_1\|_2)+(1-\alpha)(R(u_2)+L\|\delta_2-u_2\|_2)$$
$$= \alpha\bar{R}(\delta_1) + (1-\alpha)\bar{R}(\delta_2),$$

and taking the $\inf$ over Δ on the left hand-side we obtain the convexity of $\bar{R}$. □

3.8.2 Conditional expectation: notation and abuse of notations

In the analysis of the algorithm, we will consider conditional expectation by fixing some of the random variables to specific values. To simplify notations, we will use the following notation: if X and Y are random variables and g is a measurable function, then we denote $E_X[g(X,Y)]$ the expectation with respect to X, that is:

$$E_X[g(X,Y)] = E[g(X',Y)|Y],$$

where X' is a random variable independent of Y that has the same law as X. This is equivalent to the random variable obtained by fixing Y and taking the expectation with respect to the law of X only.

In line 0, we have assumed that the decision maker is able to compute $\mathbb{E}[f_t(x_t) \mid c_{tk_t}]$ and $\mathbb{E}[a_t(x_t) \mid c_{tk_t}]$. While the law of c_{tk_t} is complicated (because it involves all past decision up to time t), what the decision maker needs to compute is the value $\zeta(c_{tk_t}, x, k_t)$, where $\zeta(c, x, k) = \mathbb{E}[f_t(x) \mid c_{tk} = c]$. Those two quantities are equal because users are *i.i.d.* and k_t and f_t are independent (conditioned on z_t and the history).

This manipulation is properly stated by lemma 3.11. Indeed by denoting $\mathcal{H}_{t-1}$ the whole history up to time $t-1$, and applying this Lemma: for x_t the decision generated by the algorithm which is $(\mathcal{H}_{t-1}, k_t, c_{tk_t})$ measurable, we have that

$$\mathbb{E}[f_t(x_t) \mid c_{tk_t}, \mathcal{H}_{t-1}, k_t] = \zeta(c_{tk_t}, x_t, k_t) = \mathbb{E}_{f_t}[f_t(x_t) \mid c_{tk_t}]. \tag{3.15}$$

It is identical for $a_t(x_t)$: for $\zeta'(c, x, k) = \mathbb{E}[a_t(x) \mid c_{tk} = c]$ we have the following equality

$$\mathbb{E}[a_t(x_t) \mid c_{tk_t}, \mathcal{H}_{t-1}, k_t] = \zeta'(c_{tk_t}, x_t, k_t) = \mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}]. \tag{3.16}$$

The following Lemma is a modification of the so-called "Freezing Lemma" (see Lemma 8 of [CC20]) which helps simplifying conditional expectations by allowing you to "fix" the random variable if it is measurable with respect to the conditioning of the expectation:

Lemma 3.11. *Let X,Y and Z be random variables with (X, Z) independent of Y, and g some bounded measurable function. Let us denote by $\sigma(X)$ the sigma-algebra generated by the random variable X. Then*

$$\mathbb{E}[g(X, Y, Z) \mid \sigma(Y, Z)] = \mathbb{E}_X[g(X, Y, Z) \mid \sigma(Z)].$$

Proof. Let $V(Y, Z) = \mathbb{E}_X[g(X, Y, Z) \mid \sigma(Z)]$. To show that W is the conditional expectation of $g(X, Y, Z)$ given $\sigma(Y, Z)$, we need to show that V is $\sigma(Y, Z)$ measurable and that for any bounded W that is $\sigma(Y, Z)$ measurable we have

$$\mathbb{E}[VW] = \mathbb{E}[g(X, Y, Z)W].$$

We denote by μ the distribution of (X, Z) and ν the distribution of Y. By the Doob-Dynkin Lemma, because W is $\sigma(Y, Z)$ measurable, we have that $W = h(Y, Z)$ where h is measurable. First, by independence of Y and (X, Z), we have that the distribution of (X, Y, Z) is $d\nu \otimes d\mu$. Thus

$$\mathbb{E}[g(X, Y, Z)W] = \int \int g(x, y, z)h(y, z)d\nu(x, z)d\nu(y).$$

Now let us compute the other expectation using Fubini theorem as both g and h are bounded:

$$\begin{aligned}
\mathbb{E}[VW] &= \int \int V(y, z)h(y, z)d\mu(z)d\nu(y) \\
&= \int d\nu(y) \left(\int V(y, z)h(y, z)d\mu(z) \right) \\
&= \int d\nu(y) \left(\mathbb{E}_Z[V(y, Z)h(y, Z)] \right).
\end{aligned}$$

The random variable $h(y, Z)$ is $\sigma(Z)$ measurable, thus because $V(y, Z)$ is a conditional expectation given $\sigma(Z)$, we have

$$\mathbb{E}_Z[V(y, Z)h(y, Z)] = \mathbb{E}_{X,Z}[g(X, y, Z)h(y, Z)] = \int g(x, y, z)h(y, z)d\mu(x, z).$$

Using Fubini again we obtain

$$\begin{aligned}\mathbb{E}[VW] &= \int d\nu(y)\left(\int g(x,y,z)h(y,z)d\mu(x,z)\right)\\ &= \int\int g(x,y,z)h(y,z)d\mu(x,z)d\nu(y)\\ &= \mathbb{E}[g(X,Y,Z)W],\end{aligned}$$

which proves that this is indeed the conditional expectation, as the lemma stated. □

3.8.3 Boundedness of λ_t — proof of lemma 3.3

We next show the proof of the boundedness of the λ_t for the sequence generated by our algorithm 1.

Lemma. *Let $\lambda_0 \in \mathbb{R}^d$, $\eta > 0$, and an arbitrary sequences $(\delta_1, \delta_2, \dots) \in \Delta^{\mathbb{N}}$. We define recursively λ_t as follows:*

$$\begin{aligned}\gamma_t &= \underset{\gamma\in\Delta_{\delta_t}}{\arg\max}\{\langle\lambda_t,\gamma\rangle - \bar{R}(\gamma)\},\\ \lambda_{t+1} &= \lambda_t + \eta(\delta_t - \gamma_t).\end{aligned}$$

Then for $\Lambda = \cup_{\delta\in\bar{\Delta}}\partial\bar{R}(\delta)$ we have the following upper bound:

$$\|\lambda_t\|_2 \leqslant L + 2\eta\,\mathrm{Diam}(\Delta) + \mathrm{Dist}(\lambda_0, \Lambda + B(0, 2\eta\,\mathrm{Diam}(\Delta))),$$

where $B(0, 2\eta\,\mathrm{Diam}(\Delta))$ is the ball centered in 0 and of diameter $2\eta\,\mathrm{Diam}(\Delta)$, with the diameter and the distance taken with respect to the euclidean L_2 norm.

Proof. Recall that $\bar{\Delta} = \cup_{\delta\in\Delta}\Delta_\delta$, with $\Delta_\delta = \{\gamma : \|\gamma-\delta\|_2 \leq \mathrm{Diam}(\Delta)\}$. Hence, for any $\delta \in \Delta$ and $\gamma \in \bar{\Delta}$, we have $\|\gamma - \delta\|_2 \leq 2\,\mathrm{Diam}(\Delta)$. Let $\Lambda = \cup_{\delta\in\bar{\Delta}}\partial\bar{R}(\delta)$, and $\mathcal{C} = \mathrm{Conv}(\Lambda) + B(0, \eta 2\,\mathrm{Diam}(\Delta))$. Using that $\bar{R}$ is L-Lipschitz, we have that Λ is bounded by L for the norm $\|\cdot\|_2$, and so does $\mathrm{Conv}(\Lambda)$. Thus $\mathcal{C}$ is bounded by $L + 2\eta\,\mathrm{Diam}(\Delta)$ for $\|\cdot\|_2$. This implies that, to prove the theorem, it suffices to show that $\mathrm{Dist}(\lambda_{t+1}, \mathcal{C}) \leqslant \mathrm{Dist}(\lambda_t, \mathcal{C})$ is true for all t. Indeed, this would imply that

$$\|\lambda_t\|_2 \leq \sup_{c\in\mathcal{C}}\|c\|_2 + \mathrm{Dist}(\lambda_t, \mathcal{C}) \leq L + 2\eta\,\mathrm{Diam}\,\Delta + \mathrm{Dist}(\lambda_0, \mathcal{C}).$$

In the remainder of the proof, we show that $\mathrm{Dist}(\lambda_{t+1}, \mathcal{C}) \leqslant \mathrm{Dist}(\lambda_t, \mathcal{C})$. First, let us remark that γ_t is the solution of a convex optimization problem, wit a constraint set Δ_{δ_t} that can be rewritten as a single inequality : $h_t(\gamma) \leqslant 0$, with $h_t(\gamma) = \|\gamma - \delta_t\|_2^2 - \mathrm{Diam}(\Delta)^2$. The derivative of this function $h_t()$ is$\nabla h_t(\gamma) = 2(\gamma - \delta_t)$. As the point δ_t is in the interior (assuming that $\mathrm{Diam}(\Delta) > 0$) of Δ_{δ_t}, Slater's conditions hold, and we can apply the KKT conditions at the optimal point (γ_t). This implies that there exists $\mu \geq 0$ such that $0 \in -\lambda_t + \partial \bar{R}(\gamma_t) + \mu 2(\gamma_t - \delta_t)$. This implies that there exists $\lambda_{\delta_t} \in \partial \bar{R}(\gamma_t)$ such that

$$\lambda_{\delta_t} - \lambda_t = 2\mu(\delta_t - \gamma_t).$$

The norm of the left-hand side of the above equation must be equal to the norm of its right-hand side, which means that $\|\lambda_{\delta_t} - \lambda_t\|_2 = 2\mu\|\delta_t - \gamma_t\|_2$. This determines the value of μ. Let $\alpha_t = \eta\|\delta_t - \gamma_t\|_2 / \|\lambda_{\delta_t} - \lambda_t\|_2$, the value λ_{t+1} is therefore equal to:

$$\lambda_{t+1} = \lambda_t + \eta(\delta_t - \gamma_t) = \lambda_t + \alpha_t(\lambda_{\delta_t} - \lambda_t) = (1 - \alpha_t)\lambda_t + \alpha_t \lambda_{\delta_t}.$$

We now consider multiple cases:

Case 1. If $\alpha_t > 1$, then we rewrite λ_{t+1} as:

$$\begin{aligned}
\lambda_{t+1} &= \lambda_{\delta_t} + (1 - \alpha_t)(\lambda_t - \lambda_{\delta_t}) \\
&= \lambda_{\delta_t} + \eta(\|\delta_t - \gamma_t\|_2 - \|\lambda_{\delta_t} - \lambda_t\|_2) \frac{\lambda_t - \lambda_{\delta_t}}{\|\lambda_t - \lambda_{\delta_t}\|_2}.
\end{aligned}$$

By assumption, the $\lambda_{\delta_t} \in \partial \bar{R}(\gamma_t) \subset \Lambda$. Moreover, the norm of the second term is bounded by $2\eta \,\mathrm{Diam}(\Delta)$ because

$$|\|\delta_t - \gamma_t\|_2 - \|\lambda_{\delta_t} - \lambda_t\|_2| = \|\delta_t - \gamma_t\|_2 - \|\lambda_{\delta_t} - \lambda_t\|_2 \leqslant 2\,\mathrm{Diam}(\Delta).$$

This implies that $\lambda_{t+1} \in \mathrm{Conv}(\Lambda) + B(0, 2\eta\,\mathrm{Diam}(\Delta)) = \mathcal{C}$, and therefore that $\mathrm{Dist}(\lambda_{t+1}, \mathcal{C}) = 0 \leq \mathrm{Dist}(\lambda_t, \mathcal{C})$.

Case 2. If $\alpha_t \leq 1$ and $\lambda_t \in \mathcal{C}$. In this case, λ_{t+1} is a convex combination of λ_{δ_t} and λ_t. As, $\lambda_{\delta_t} \in \Lambda \subset \mathcal{C}$ and $\lambda_t \in \mathcal{C}$, and because $\mathcal{C}$ is convex, we have that $\lambda_{t+1} \in \mathcal{C}$.

Case 3. If $\alpha_t \leq 1$ and $\lambda_t \notin \mathcal{C}$, we define $\pi_t = \lambda_t$ the projection of λ_t on $\mathcal{C}$. Let $y_t = (1 - \alpha_t)\pi_t + \alpha_t \lambda_{\delta_t}$. Because $\mathcal{C}$ is convex, $\alpha_t \in [0, 1]$, $\lambda_{\delta_t} \in \mathcal{C}$, and $\pi_t \in \mathcal{C}$, we have $y_t \in \mathcal{C}$. This implies that

$$\begin{aligned}
\mathrm{Dist}(\lambda_{t+1}, \mathcal{C}) &\leqslant \|\lambda_{t+1} - y_t\|_2 \\
&= (1 - \alpha_t)\|\lambda_t - \pi_t\|_2 \\
&= (1 - \alpha_t)\,\mathrm{Dist}(\lambda_t, \mathcal{C}) \\
&\leqslant \mathrm{Dist}(\lambda_t, \mathcal{C}).
\end{aligned}$$

This shows that for all cases, $\mathrm{Dist}(\lambda_{t+1}, \mathcal{C}) \leqslant \mathrm{Dist}(\lambda_t, \mathcal{C})$, which concludes the proof. □

Using this previous Lemma, we can simply bound the virtual reward functions.

Lemma 3.12. *For $\lambda_0 \in \partial R(0)$, and for λ_t generated according to algorithm 1, the virtual reward function*

$$\varphi(\lambda_t, c_{tk}, k) = \max_{x \in \mathcal{X}} \mathbb{E}[f_t(x) - \langle \lambda_t, a_t(x) \rangle \mid c_{tk}] - p_k,$$

is bounded in absolute value by

$$m := \bar{f} + L + 2\eta \,\mathrm{Diam}(\Delta) + \max_k |p_k|.$$

Proof. Because $a_t(x)$ used for the update of λ_t is in Δ for all t, we can apply lemma 3.3 to bound the trajectory of the $(\lambda_t)_{t \in [T]}$ generated by the algorithm. Using a triangle inequality first, then Cauchy-Schwartz inequality, and finally the boundedness assumptions on f_t and a_t with the Lemma, we have

$$\begin{aligned}
&|\max_{x \in \mathcal{X}} \mathbb{E}[f_t(x) - \langle \lambda_t, a_t(x) \rangle \mid c_{tk}] - p_k| \\
&\leqslant |\max_{x \in \mathcal{X}} \mathbb{E}[f_t(x) \mid c_{tk}]| + |\max_{x \in \mathcal{X}} \mathbb{E}[\langle \lambda_t, a_t(x) \rangle \mid c_{tk}]| + |p_k| \\
&\leqslant |\max_{x \in \mathcal{X}} \mathbb{E}[f_t(x) \mid c_{tk}]| + |\max_{x \in \mathcal{X}} \mathbb{E}[\|\lambda_t\|_2 \|a_t(x)\|_2 \mid c_{tk}]| + |p_k| \\
&\leqslant \bar{f} + L + 2\eta \,\mathrm{Diam}(\Delta) + \max_k |p_k|.
\end{aligned}$$

□

Lemma 3.13. *For $\lambda_0 \in \partial R(a)$ (for some $a \in \mathcal{A}$), and for λ_t generated according to algorithm 1, if the quantity $a_t(x)/x$ is always well defined and in Δ for all a_t and all x, then the virtual reward function*

$$\tilde{\varphi}(\lambda_t, c_{tk}, k) = \max_{x \in \mathcal{X}} \left(\mathbb{E}[f_t(x) - \langle \lambda_t, a_t(x) \rangle \mid c_{tk}] + R^*(\lambda_t) \right) - p_k,$$

is bounded in absolute value by

$$\tilde{m} := \bar{f} + L + 2\eta \,\mathrm{Diam}(\Delta) + \max\{R^*(\lambda) \mid \|\lambda\|_2 \leqslant L + 2\eta \,\mathrm{Diam}(\Delta)\} + \max_k |p_k|.$$

Proof. The proof of this Lemma is identical to the previous one, except that we also need to bound R^*. As a convex conjugate, R^* is convex. By Corollary 10.1.1 of

[Roc70], every convex function from $\mathbb{R}^d$ to $\mathbb{R}$ is continuous. By Lemma 1 of [Cel+22], we know that the domain of R^* is indeed $\mathbb{R}^d$, hence it is continuous.

Now let us look at the modified dual update $\lambda_{t+1} = \lambda_t + \eta(\gamma_t x_t - a_t(x_t))$. We can factorize by x_t to recover $\lambda_{t+1} = \lambda_t + x_t\eta(\gamma_t - a_t(x_t)/x_t)$. Because of our hypothesis, $a_t(x_t)/x_t \in \Delta$, therefore we can apply lemma 3.3 with $\eta_t = \eta x_t$, which is upper bounded by η. Hence the λ_t are all in $\{\lambda \in \mathbb{R}^d \mid \|\lambda\|_2 \leqslant L + 2\eta\,\mathrm{Diam}(\Delta)\}$ which is compact. By continuity of R^*, it is continuous over this compact set hence bounded, and the virtual reward is bounded by $\tilde{m}$ overall. □

Remark. *We made the technical assumption that $a_t(x)/x \in \Delta$. It is clear that when a_t is linear, such as $a_t x$, then $a_t x/x = a_t \in \mathcal{A} \subset \Delta$ and the condition is satisfied. Regardless, if $a_t(x)/x$ is always well defined and bounded, we can simply increase the size of Δ, so that this quantity always remain in Δ. Of course, this will degrade the performance of the algorithm as it depends on* $\mathrm{Diam}(\Delta)$.

3.8.4 Expected difference between penalty and penalty of expectation — Proof of lemma 3.4

We next prove the following lemma, which tells us that observing the conditional expectation of the $a_t(x_t)$ is enough to observe the expectation of the penalty R:

Lemma. *For $(x_1, \dots, x_T)$ and $(k_1, \dots, k_T)$ generated according to algorithm 1, we have the following upper bound:*

$$|\mathbb{E}[R(\frac{1}{T}\sum_{t=1}^{T} a_t x_t) - R(\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}])]| \leqslant 2L\sqrt{\frac{d}{T}}.$$

Proof. Let us denote by $\delta_t = \mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}]$ the expectation of $a_t(x_t)$ with x_t fixed. Using first the triangle inequality for the expectation, and then the Lipschitzness of R we have that

$$\begin{aligned}
|\mathbb{E}[R(\frac{1}{T}\sum_{t=1}^{T} a_t(x_t)) - R(\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}])]| &\leqslant \mathbb{E}[|R(\frac{1}{T}\sum_{t=1}^{T} a_t(x_t)) - R(\frac{1}{T}\sum_{t=1}^{T}\delta_t)|] \\
&\leqslant \frac{L}{T}\mathbb{E}[\|\sum_{t=1}^{T}(a_t(x_t) - \delta_t)\|_2] \\
&\leqslant \frac{L}{T}\sqrt{\mathbb{E}[\|\sum_{t=1}^{T}(a_t(x_t) - \delta_t)\|_2^2]}, \quad (3.17)
\end{aligned}$$

where we used Jensen's inequality with $x \mapsto \sqrt{x}$ for the last inequality.

We denote by $\delta_{t,l}$ the l-th coordinate of δ_t for $l \in [d]$, and similarly for $a_{t,l}$. We have:

$$\mathbb{E}[\|\sum_{t=1}^{T}(a_t(x_t) - \delta_t)\|_2^2] = \sum_{l=1}^{d}\mathbb{E}[\left(\sum_{t=1}^{T}(a_{t,l}(x_t) - \delta_{t,l})\right)^2]$$

By eq. (3.16), the sequence $a_{t,l}(x_t) - \delta_{t,l}$ is a Martingale difference sequence for the filtration $\sigma((a_\tau, c_{\tau k_\tau}, k_\tau)_{\tau\in[t-1]}, k_t, c_{tk_t})$. Hence, using that the expectation is 0 by the law of total expectation, we obtain that

$$\mathbb{E}\left[\left(\sum_{t=1}^{T}(a_{t,l}(x_t) - \delta_{t,l})\right)^2\right] = \mathrm{Var}[\sum_{t=1}^{T}(a_{t,l}(x_t)-\delta_{t,l})] = \sum_{t=1}^{T}\mathrm{Var}[a_{t,l}(x_t)-\delta_{t,l}]. \quad (3.18)$$

The terms $a_{t,l}(x_t) - \delta_{t,l}$ are bounded in $[-2, 2]$ by the assumption that $\|a_t(x)\|_2 \leqslant 1$, therefore we can bound each of these variances by 4, for a total bound of $4T$. Plugging this into eq. (3.17) concludes the proof. □

3.9 Proof of the main theorem 3.2

The proof builds upon the proof used in the full information case (a_t and f_t are given) from [BLM21; Cel+22].

In order to bound the regret, we upper bound the performance of OPT, and lower bound the performance of ALG. We proceed by first upper-bounding the optimum.

Recall that the convex conjugate R^* of R (which we consider as a function from $\mathbb{R}^d \to \mathbb{R}$ that is $+\infty$ outside of Δ) is a function that associates to a dual parameter λ the quantity $R^*(\lambda)$:

$$R^*(\lambda) = \max_{\gamma\in\mathbb{R}^d}\{\langle\gamma, \lambda\rangle - R(\gamma)\} = \max_{\gamma\in\Delta}\{\langle\gamma, \lambda\rangle - R(\gamma)\}.$$

The Fenchel-Moreau theorem, tells us that for a proper convex function, the biconjugate is equal to the original function: $R^{**} = R$.

Recall that φ is defined in eq. (3.5). We define the dual vector function $\mathcal{D}(\lambda) = (\mathcal{D}(\lambda, 1), \dots, \mathcal{D}(\lambda, K))$ which is a vector representing the value of the dual conjugate problem. For coordinate $k \in [K]$, it is defined as

$$\mathcal{D}(\lambda, k) = \mathbb{E}[\varphi(\lambda, c_{tk}, k)] + R^*(\lambda).$$

We denote the unit simplex of dimension K by $\mathcal{P}_K = \{\pi \in \mathbb{R}^K \mid \pi_k \geq 0, \sum_{k=1}^K \pi_k = 1\}$.

Lemma 3.14. *We have the following upper-bound for the offline optimum:*

$$\mathrm{OPT} \leqslant T \sup_{\pi \in \mathcal{P}_K} \inf_{\lambda \in \mathbb{R}^d} \langle \pi, \mathcal{D}(\lambda) \rangle,$$

Proof. We upper bound the performance of OPT through a Lagrangian relaxation using the convex conjugate of R, as was done similarly in [Jen+16].

Let $(k_1, \ldots, k_T) \in [K]^T$ be the (deterministic) sequence of sources selected, and $(x_1, \ldots, x_T) \in \mathcal{X}^T$ the offline allocation. These x_t are all $\sigma(c_{1k_1}, \ldots, c_{Tk_T})$ measurable. We define $\tilde{f}_t(x) = \mathbb{E}_{f_t}[f_t(x) \mid c_{tk_t}]$ and $\delta_t = \mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}]$.

Using Jensen's inequality for R convex we have

$$\begin{aligned}
&\mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid c_{1k_1}, \ldots, c_{Tk_T}] \\
&= \sum_{t=1}^T \mathbb{E}[f_t(x_t) - p_{k_t} \mid c_{1k_1}, \ldots, c_{Tk_T}] - T\mathbb{E}[R(\frac{1}{T}\sum_{t=1}^T a_t(x_t)) \mid c_{1k_1}, \ldots, c_{Tk_T}] \\
&\leqslant \sum_{t=1}^T \mathbb{E}[f_t(x_t) - p_{k_t} \mid c_{1k_1}, \ldots, c_{Tk_T}] - TR(\mathbb{E}[\frac{1}{T}\sum_{t=1}^T a_t(x_t) \mid c_{1k_1}, \ldots, c_{Tk_T}]).
\end{aligned}$$

Because the x_t are $\sigma(c_{1k1}, \ldots, c_{Tk_T})$ measurable, and by independence of the $(f_t, a_t, c_{t1}, \ldots, c_{tK})$ we have that $\mathbb{E}[a_t(x_t) \mid \sigma(c_{1k_1}, \ldots, c_{tk_T})] = \mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}] = \delta_t$. With the same argument we obtain that $\mathbb{E}[f_t(x_t) \mid \sigma(c_{1k_1}, \ldots, c_{tk_t})] = \tilde{f}_t(x_t)$. Therefore

$$\mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid c_{1k_1}, \ldots, c_{Tk_T}] \leqslant \sum_{t=1}^T \tilde{f}_t(x_t) - p_{k_t} - TR(\frac{1}{T}\sum_{t=1}^T \delta_t). \tag{3.19}$$

We define the function $\mathcal{L} : (\boldsymbol{k}, \mathbf{c}, \mathbf{x}, \lambda) \mapsto \sum_{t=1}^T \tilde{f}_t(x_t) - p_{k_t} - \langle \lambda, \delta_t \rangle + TR^*(\lambda)$.

We now show that this function is always greater than the conditional expectation of $\mathcal{U}$:

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{k}, \mathbf{c}, \mathbf{x}, \lambda) &\geqslant \inf_{\lambda \in \mathbb{R}^d} \mathcal{L}(\boldsymbol{k}, \mathbf{c}, \mathbf{x}, \lambda) \\
&= \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - T \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, \frac{1}{T} \sum_{t=1}^{T} \delta_t \rangle - R^*(\lambda)\} \\
&= \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - TR^{**}(\frac{1}{T} \sum_{t=1}^{T} \delta_t) \\
&= \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - TR(\frac{1}{T} \sum_{t=1}^{T} \delta_t) \\
&\geqslant \mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid \mathbf{c}_{\boldsymbol{k}}], \qquad (3.20)
\end{aligned}
$$

the last equality results from applying Fenchel-Moreau theorem, and the last inequality is obtained by using eq. (3.19).

Let us now compute the max in $(x_1, \ldots, x_T)$ of $\mathcal{L}$. Looking at the terms inside the sum, the only possible dependency of $\tilde{f}_t(x_t) - p_{k_t} - \langle \lambda, \delta_t \rangle$ to other contexts for $t' \neq t$, is possible through the allocation x_t, which can depend on the other contexts. Therefore the maximum of the sum is the sum of the maximums. This yields that

$$
\begin{aligned}
\max_{(x_1, \ldots, x_T)} \mathcal{L}(\boldsymbol{k}, \mathbf{c}, \mathbf{x}, \lambda) &= \sum_{t=1}^{T} \max_{x_t \in \mathcal{X}} \mathbb{E}_{f_t, a_t}[f_t(x_t) - p_{k_t} - \langle \lambda, a_t(x_t) \rangle \mid c_{tk_t}] + TR^*(\lambda) \\
&= \sum_{t=1}^{T} (\varphi(\lambda, c_{tk_t}, k_t) + R^*(\lambda)).
\end{aligned}
$$

Thus by taking the max in $(x_1, \ldots, x_T)$ on both sides of eq. (3.20) and then taking the expectation (which is over the contexts as everything is measurable with respect to them), we obtain

$$
\mathbb{E}[\max_x \mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid c_{1k_1}, ., c_{Tk_T}]] \leqslant \sum_{t=1}^{T} \mathbb{E}[\varphi(\lambda, c_{tk_t}, k_t) + R^*(\lambda)] = \sum_{t=1}^{T} \mathcal{D}(\lambda, k_t). \quad (3.21)
$$

What is nice here, is that by looking at the dual problem which was originally to address the non-separability in x, we also have the separability in k_t.

We denote by $T_k = \sum_{t=1}^{T} \mathbb{1}[k_t = k]$ the number of times source k is selected, and by $\widehat{\pi} = (T_1/T, \ldots, T_K/T) \in \mathcal{P}_K$ the selection proportions of each source k over the

T rounds. This vector of proportions can be seen as the empirical distribution of selected sources. We can rewrite in the following way the dual objective

$$\sum_{t=1}^{T} \mathcal{D}(\lambda, k_t) = \sum_{k=1}^{K} T_k \mathcal{D}(\lambda, k) = T \sum_{k=1}^{K} \widehat{\pi}_k \mathcal{D}(\lambda, k) = T \langle \widehat{\pi}, \mathcal{D}(\lambda) \rangle.$$

Finally, using the previous equation with eq. (3.21), taking first the $\inf$ in λ and then the $\max$ in k_t we can conclude:

$$\mathrm{OPT} \leqslant T \max_{k_1,\ldots,k_T} \inf_{\lambda \in \mathbb{R}^d} \langle \widehat{\pi}, \mathcal{D}(\lambda) \rangle \leqslant T \sup_{\pi \in \mathcal{P}_K} \inf_{\lambda \in \mathbb{R}^d} \langle \pi, \mathcal{D}(\lambda) \rangle. \qquad \square$$

Remark. *Note that the function $\mathcal{D}(\lambda)$ is convex and lower semi-continuous (l.s.c.) (as the supremum of a family of l.s.c. functions). In addition, $\langle \pi, \mathcal{D}(\lambda) \rangle$ is concave and u.s.c. in π by linearity, and $\mathcal{P}_K$ is compact convex. Hence Sion's minimax theorem can be applied, and the $\sup$ in π is actually a $\max$.*

We can now proceed to the rest of the proof of the main theorem 3.2:

Theorem 3.15. *For $\eta = L/(2\,\mathrm{Diam}(\Delta)\sqrt{T})$, $m = \bar{f} + L + \max_k |p_k| + 2\eta\,\mathrm{Diam}(\Delta)$, $\rho = \sqrt{\log(K)/(TKm^2)}$, and $\lambda_0 \in \partial R(0)$, algorithm algorithm 1 has the following regret upper bound:*

$$\mathrm{Reg} \leqslant 2((L + \bar{f} + \max_k |p_k|)\sqrt{K \log(K)} + L\sqrt{d} + \mathrm{Diam}(\Delta))\sqrt{T} + 2L\sqrt{K \log(K)}$$

Proof. We first lower bound the performance of the algorithm. Then we simply need to apply lemma 3.14 to get the regret bound. Let $\mathcal{H}_t$ be the natural filtration generated by the $(c_{\tau k_\tau}, k_\tau)_{\tau \in [t]}$. Let x_t, k_t, λ_t and γ_t defined as in 1. We denote by $\delta_t = \mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}]$. The quantity δ_t is in Δ because of the expectation and by convexity of Δ. Note that here these are all random variables, with γ_t and λ_t being $\mathcal{H}_{t-1}$ measurable and x_t being $\mathcal{H}_t$ measurable.

To first get a broad picture on how we are going to lower bound the utility of the algorithm, let us decompose it into different parts. It can be rewritten as

$$\begin{aligned}
\mathbb{E}[\sum_{t=1}^{T} f_t(x_t) - p_{k_t} - TR(\frac{1}{T}\sum_{t=1}^{T} a_t(x_t))] =& \mathbb{E}\Bigg[\left(\sum_{t=1}^{T} f_t(x_t) - p_{k_t} - \bar{R}(\gamma_t)\right) && (3.22)\\
&+ \left(\sum_{t=1}^{T} \bar{R}(\gamma_t) - TR(\frac{1}{T}\sum_{t=1}^{T} \delta_t)\right) && (3.23)\\
&+ \left(TR(\frac{1}{T}\sum_{t=1}^{T} \delta_t) - TR(\frac{1}{T}\sum_{t=1}^{T} a_t(x_t))\right)\Bigg]. && (3.24)
\end{aligned}$$

The first line (eq. (3.22)) corresponds to some adjusted separable rewards, and we bound it by first reducing it to the virtual rewards φ, and then using bandits algorithms techniques. The second line (eq. (3.23)) corresponds to the difference between the unfairness of the adjusted rewards, and the true unfairness, which is not big because of the gradient descent on λ_t. Finally the last line (eq. (3.24)) corresponds to the difference between the penalty of the expected allocation and the expected penalty of the allocation, which we have already shown in lemma 3.4 is bounded by $\mathcal{O}(\sqrt{T})$.

Reduction of eq. (3.22) to a bandit problem with φ Using first the tower property of the conditional expectation, and then *eq.* (3.15), we have that

$$\mathbb{E}[f_t(x_t)] = \mathbb{E}\left[\mathbb{E}[f_t(x_t) \mid \mathcal{H}_t]\right] = \mathbb{E}\left[\mathbb{E}_{f_t}[f_t(x_t) \mid c_{tk_t}]\right].$$

We add and remove $\langle \lambda_t, \delta_t \rangle$, and use that x_t is the allocation that yields the virtual reward function:

$$\begin{aligned}\mathbb{E}[f_t(x_t) - p_{k_t}] &= \mathbb{E}\left[\mathbb{E}_{f_t}[f_t(x_t) \mid c_{tk_t}] - p_{k_t} - \langle \lambda_t, \delta_t \rangle + \langle \lambda_t, \delta_t \rangle\right] \\ &= \mathbb{E}[\varphi(\lambda_t, c_{tk_t}, k_t) + \langle \lambda_t, \delta_t \rangle].\end{aligned}$$

Using the definition of γ_t as a maximizer, we have

$$\begin{aligned}-\bar{R}(\gamma_t) &= -\bar{R}(\gamma_t) + \langle \lambda_t, \gamma_t \rangle - \langle \lambda_t, \gamma_t \rangle \\ &= \max_{\gamma \in \Delta_{\delta_t}} \{\langle \lambda_t, \gamma \rangle - \bar{R}(\gamma)\} - \langle \lambda_t, \gamma_t \rangle.\end{aligned}$$

Because $\Delta \subset \Delta_{\delta_t}$ and, because $\bar{R}$ and R are equal over Δ we recover a lower bound with R^*:

$$\begin{aligned}\max_{\gamma \in \Delta_{\delta_t}} \{\langle \lambda_t, \gamma \rangle - \bar{R}(\gamma)\} - \langle \lambda_t, \gamma_t \rangle &\geqslant \max_{\gamma \in \Delta}\{\langle \lambda_t, \gamma \rangle - \bar{R}(\gamma)\} - \langle \lambda_t, \gamma_t \rangle \\ &= \max_{\gamma \in \Delta}\{\langle \lambda_t, \gamma \rangle - R(\gamma)\} - \langle \lambda_t, \gamma_t \rangle \\ &= R^*(\lambda_t) - \langle \lambda_t, \gamma_t \rangle.\end{aligned}$$

Thus

$$\mathbb{E}[f_t(x_t) - p_{k_t} - \bar{R}(\gamma_t)] \geqslant \mathbb{E}[\varphi(\lambda_t, c_{tk_t}, k_t) + R^*(\lambda_t) + \langle \lambda_t, \delta_t - \gamma_t \rangle]. \tag{3.25}$$

Application of Bandits Algorithm Let π_t be the distribution generated according to the algorithm. We actually have that $\mathbb{E}[\varphi(\lambda_t, c_{tk_t}) + R^*(\lambda_t)]$ is just (through independence, measurability, and total expectation arguments) $\mathbb{E}[\langle \pi_t, \mathcal{D}(\lambda_t)\rangle]$. Hence the main idea is to apply Online Convex Optimization theorems to the linear rewards $\mathcal{D}(\lambda_t)$ with the decision variable π_t (see [Haz22], [Ora19] or [LS20a] for introductions to these types of problems). Here we don't have access either to $\mathcal{D}(\lambda_t)$, nor even to $\mathcal{D}(\lambda_t, k_t)$: we are in an adversarial bandit setting regarding the λ_t, and also in a stochastic setting regarding the expectation in $\mathcal{D}$ taken with respect to c_t. Hence we use an unbiased estimator of the gradient $\mathcal{D}(\lambda_t)$ for the linear gain, which we will describe further down.

Notice that it is crucial to be able to deal with any sequence $\mathcal{D}(\lambda_t)$, this is because these are neither independent nor related to martingales, as the λ_t are generated from a quite complicated Markov Chain with states $(\lambda_t, c_t, \pi_t, k_t)$. Hence the outside expectation does not help us much, and we will face as losses (or here rewards) the random arbitrary sequence of the $\mathcal{D}(\lambda_t)$. We don't care about $R^*(\lambda_t)$ as it is $\mathcal{H}_{t-1}$ measurable.

Here the different arms of a bandit problem correspond to different sources. Using the tower rule, we have

$$\mathbb{E}\left[\sum_{t=1}^{T} \varphi(\lambda_t, c_{tk_t}, k_t)\right] = \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{E}[\varphi(\lambda_t, c_{tk_t}, k_t) \mid \mathcal{H}_{t-1}]\right].$$

To use results for adversarial bandits, the bandits rewards need to be bounded. By lemma 3.12, the virtual reward φ, and consequently its conditional expectation, is bounded by $m := \bar{f} + L + 2\eta\,\mathrm{Diam}(\Delta) + \max_k |p_k|$. Therefore $\mathbb{E}[\varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}]/m \in [-1, 1]$.

We now construct an unbiased estimator of $\mathbb{E}[\varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}]$, which is the following importance weighted estimator:

$$\widehat{\varphi}(\lambda_t, c_{tk}, k) = m - \mathbb{1}[k_t = k]\frac{(m - \varphi(\lambda_t, c_{tk}, k))}{\pi_{tk}}.$$

This estimator is unbiased for the conditional expectation given $\mathcal{H}_{t-1}$. Indeed because π_{tk} is $\mathcal{H}_{t-1}$ measurable, and because $\mathbb{1}[k_t = k]$ and $\varphi(\lambda_t, c_{tk}, k)$ are independent conditionally on $\mathcal{H}_{t-1}$ we have:

$$\begin{aligned}
\mathbb{E}[\widehat{\varphi}(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}] &= m - \frac{1}{\pi_{tk}}\mathbb{E}[(m - \varphi(\lambda_t, c_{tk}, k))\mathbb{1}[k_t = k] \mid \mathcal{H}_{t-1}] \\
&= m - \frac{1}{\pi_{tk}}\mathbb{E}[m - \varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}]\mathbb{E}[\mathbb{1}[k_t = k] \mid \mathcal{H}_{t-1}] \\
&= m - \mathbb{E}[m - \varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}] \\
&= \mathbb{E}[\varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}],
\end{aligned}$$

which is exactly the reward needed.

We can now use the following Theorem 11.2 from [LS20a]:

Theorem (EXP3 Regret Bound). *Let $X \in [0,1]^{T\times K}$ be the rewards of an adversarial bandit, then running EXP3 with with learning rate $\rho = \sqrt{2\log(K)/(TK)}$ and denoting k_t the arm chosen at time t, achieves the following regret bound:*

$$\max_{k\in K}\sum_{t=1}^{T} X_{t,k} - \mathbb{E}[\sum_{t=1}^{T} X_{t,k_t}] \leqslant \sqrt{2TK\log(K)}$$

The fact that we can possibly have negative rewards instead of in $[0,1]$ like requested in the theorem, just changes that the loss estimator used as an intermediary is bounded by 2 instead of 1. Hence by selecting $\rho = \sqrt{\log(K)/TK}/m$, and using the importance weighted estimator $\widehat{\varphi}(\lambda_t, c_t)$ we can use the EXP3 Theorem an apply the expectation to obtain the following regret bound:

$$\mathbb{E}[\max_{k\in[K]}\sum_{t=1}^{T}\mathbb{E}[\varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}]] - \mathbb{E}[\sum_{t=1}^{T}\varphi(\lambda_t, c_{tk_t}, k_t)] \leqslant 2m\sqrt{TK\log(K)}.$$

We know that any convex combination of the sum of the rewards of multiple sources is smaller than the sum of the rewards for the best source. Let us denote by $(\pi^n)_{n\in\mathbb{N}}$ the sequence of distributions that maximizes the dual objective of lemma 3.14. We will use π^n as a convex combination, and taking the expectation over the regret bound we obtain:

$$\mathbb{E}[\sum_{t=1}^{T}\varphi(\lambda_t, c_{tk_t}, k_t)] \geqslant \mathbb{E}\left[\sum_{k=1}^{K}\pi_k^n\sum_{t=1}^{T}\mathbb{E}[\varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}]\right] - 2m\sqrt{TK\log(K)}.$$

Moreover using that c_{tk} is independent of $\mathcal{H}_{t-1}$ and that λ_t is $\mathcal{H}_{t-1}$ measurable we can deduce using the freezing Lemma that

$$\mathbb{E}[\varphi(\lambda_t, c_{tk}, k) \mid \mathcal{H}_{t-1}] + R^*(\lambda_t) = \mathbb{E}_{c_{tk}}[\varphi(\lambda_t, c_{tk}, k)] + R^*(\lambda_t) = \mathcal{D}(\lambda_t, k).$$

Hence

$$\begin{aligned}\mathbb{E}[\sum_{t=1}^{T} \varphi(\lambda_t, c_{tk_t}) + R^*(\lambda_t)] &\geqslant \mathbb{E}[\sum_{k=1}^{K} \pi_k^n \sum_{t=1}^{T} \mathcal{D}(\lambda_t, k)] - 2m\sqrt{TK\log(K)} \\ &= \mathbb{E}[\sum_{t=1}^{T} \langle \pi^n, \mathcal{D}(\lambda_t)\rangle] - 2m\sqrt{TK\log(K)} \\ &\geqslant \sum_{t=1}^{T} \inf_{\lambda\in\mathbb{R}^d} \langle \pi^n, \mathcal{D}(\lambda)\rangle - 2m\sqrt{TK\log(K)}.\end{aligned}$$

Taking the limit in $n \to \infty$ and by definition of π^n we conclude that

$$\mathbb{E}[\sum_{t=1}^{T} \varphi(\lambda_t, c_{tk_t}) + R^*(\lambda_t)] \geqslant T \sup_{\pi\in\mathcal{P}_K} \inf_{\lambda\in\mathbb{R}^d} \langle \pi, \mathcal{D}(\lambda)\rangle - 2m\sqrt{TK\log(K)}.$$

Using the above regret bound and eq. (3.25) we obtain

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{T} f_t(x_t) - p_{k_t} - \bar{R}(\gamma_t)\right] &\geqslant T \sup_{\pi\in\mathcal{P}_K} \inf_{\lambda\in\mathbb{R}^d} \langle \pi, \mathcal{D}(\lambda)\rangle - 2m\sqrt{TK\log(K)} \\ &\quad + \mathbb{E}\left[\sum_{t=1}^{T} \langle \lambda_t, \delta_t - \gamma_t\rangle\right].\end{aligned} \tag{3.26}$$

Analysis of eq. (3.23) and dual gradient descent We will now compare $TR(\sum_{t=1}^T \delta_t/T)$ and $\sum_{t=1}^T \bar{R}(\gamma_t)$. Here we consider a modified version of $\bar{R}$, which is equal to $\bar{R}$ over $\bar{\Delta}$ and is equal to $+\infty$ outside of $\bar{\Delta}$, and we will use its convex conjugate $\bar{R}^*$ (note that the convex conjugate of $\bar{R}$ without these modification would have been different !). Let

$$\widehat{\lambda} \in \arg\max_{\lambda\in\mathbb{R}^d} \{\langle \lambda, \frac{1}{T}\sum_{t=1}^{T} \delta_t\rangle - \bar{R}^*(\lambda)\}.$$

We have by the Fenchel-Moreau theorem that

$$\bar{R}(\frac{1}{T}\sum_{t=1}^{T} \delta_t) = \bar{R}^{**}(\frac{1}{T}\sum_{t=1}^{T} \delta_t) = \langle \widehat{\lambda}, \frac{1}{T}\sum_{t=1}^{T} \delta_t\rangle - \bar{R}^*(\widehat{\lambda}).$$

By definition of the convex conjugate of $\bar{R}$, for any $\gamma \in \mathbb{R}^d$, thus in particular for all γ_t, we have

$$\bar{R}^*(\widehat{\lambda}) \geqslant \langle \widehat{\lambda}, \gamma_t \rangle - \bar{R}(\gamma_t).$$

Summing for every $t \in [T]$ we obtain

$$T\bar{R}^*(\widehat{\lambda}) \geqslant \sum_{t=1}^{T} \langle \widehat{\lambda}, \gamma_t \rangle - \bar{R}(\gamma_t)$$

Hence

$$\begin{aligned} &\langle \widehat{\lambda}, \sum_{t=1}^{T} \delta_t \rangle - T\bar{R}(\frac{1}{T}\sum_{t=1}^{T} \delta_t) \geqslant \sum_{t=1}^{T} \langle \widehat{\lambda}, \gamma_t \rangle - \bar{R}(\gamma_t) \\ \Leftrightarrow \quad &\langle \widehat{\lambda}, \sum_{t=1}^{T} \delta_t - \gamma_t \rangle \geqslant T\bar{R}(\frac{1}{T}\sum_{t=1}^{T} \delta_t) - \sum_{t=1}^{T} \bar{R}(\gamma_t). \end{aligned}$$

Finally because $\bar{R}$ and R are equal over Δ, and because $\sum_{t=1}^{T} \delta_t / T \in \Delta$, we derive the following inequality:

$$\sum_{t=1}^{T} \bar{R}(\gamma_t) - TR(\frac{1}{T}\sum_{t=1}^{T} \delta_t) \geqslant \sum_{t=1}^{T} \langle \widehat{\lambda}, \gamma_t - \delta_t \rangle. \tag{3.27}$$

We now would like to track the left-hand sum evaluated at the dual parameter $\widehat{\lambda}$. To do so we will use straightforward Online Gradient Descent on the linear function $\lambda \mapsto \langle \lambda, \gamma_t - \delta_t \rangle$, with sub-gradients $\gamma_t - \delta_t$.

We will use Lemma 6.17 from [Ora19]:

Theorem (Online Gradient Descent). *For $g_t = \gamma_t - \delta_t$ a sub-gradient of a convex loss function l_t with $\|g_t\|_2 \leqslant G$, for w_t generated according to a gradient descent update with parameter η, we have that for any $u \in \mathbb{R}^d$:*

$$\sum_{t=1}^{T} l_t(w_t) - l_t(u) \leqslant \frac{\|u\|_2^2}{2\eta} + \frac{\eta}{2} T G^2.$$

This is exactly our setting, and the way we update λ_t. Therefore

$$\sum_{t=1}^{T} \langle \lambda_t, \gamma_t - \delta_t \rangle - \langle \widehat{\lambda}, \gamma_t - \delta_t \rangle \leqslant \frac{\|\widehat{\lambda}\|_2^2}{2\eta} + \frac{\eta}{2} T G^2.$$

Moreover, because $\widehat{\lambda}$ is the $\arg\max$ which yields the convex conjugate of $\bar{R}^*$, we have $\widehat{\lambda} \in \partial \bar{R}^{**}(\sum \delta_t/T) = \partial R^{**}(\sum \delta_t/T)$. Because R is L-Lipschitz for $\|\cdot\|_2$, we have $\|\widehat{\lambda}\|_2 \leqslant L$. Hence

$$\sum_{t=1}^{T} \langle \lambda_t, \gamma_t - \delta_t \rangle - \langle \widehat{\lambda}, \gamma_t - \delta_t \rangle \leqslant \frac{L^2}{2\eta} + \frac{\eta}{2} T G^2.$$

What do we lose by using $\bar{R}$ instead of R to generate γ_t in our gradient descent ? We have larger gradients bounds by a factor 2 in our case. Indeed instead of having $G \leqslant \mathrm{Diam}(\Delta)$, because γ_t can be outside of Δ, we have $\|\gamma_t - \delta_t\|_2 \leqslant 2\,\mathrm{Diam}(\Delta)$. In this case, choosing $\eta = L/(2\,\mathrm{Diam}(\Delta)\sqrt{T})$ is optimal, and yields the following regret bound:

$$\sum_{t=1}^{T} \langle \lambda_t, \gamma_t - \delta_t \rangle - \langle \widehat{\lambda}, \gamma_t - \delta_t \rangle \leqslant 2L\,\mathrm{Diam}(\Delta)\sqrt{T} \tag{3.28}$$

Using both eq. (3.27) and eq. (3.28) yield:

$$\sum_{t=1}^{T} \bar{R}(\gamma_t) - T R(\frac{1}{T} \sum_{t=1}^{T} \delta_t) \geqslant \sum_{t=1}^{T} \langle \lambda_t, \gamma_t - \delta_t \rangle - 2L\,\mathrm{Diam}(\Delta)\sqrt{T} \tag{3.29}$$

We can finally put everything together. Through the decomposition with eqs. (3.22) to (3.24), using respectively eqs. (3.26) and (3.29) and lemma 3.4 for each of those differences and summing the inequalities, the terms in $\langle \lambda_t, \gamma_t - \delta_t \rangle$ cancel out, and we obtain that

$$\mathrm{ALG} \geqslant T \sup_{\pi \in \mathcal{P}_K} \inf_{\lambda \in \mathbb{R}^d} \langle \pi, \mathcal{D}(\lambda) \rangle - 2m\sqrt{TK\log(K)} - 2L\,\mathrm{Diam}(\Delta)\sqrt{T} - 2L\sqrt{dT}.$$

By applying lemma 3.14 we conclude that

$$\begin{aligned} \mathrm{Reg} &= \mathrm{OPT} - \mathrm{ALG} \\ &\leqslant 2((L + \bar{f} + \max_k |p_k|)\sqrt{K\log(K)} + L\sqrt{d} + L\,\mathrm{Diam}(\Delta))\sqrt{T} + 2L\sqrt{K\log(K)}. \square \end{aligned}$$

3.10 The Importance of Randomizing Between Sources: OPT vs static-OPT

3.10.1 Proof of proposition 3.1

We can directly derive from the proof of lemma 3.14 and algorithm 1 the following proposition, which is a rewriting of proposition 3.1.

Proposition. *There are instances of the problem such that*

$$\lim_{T\to\infty} \frac{\text{static}-\text{OPT}}{T} < \lim_{T\to\infty} \frac{\text{OPT}}{T}.$$

Proof. What we have shown indirectly in the proof of theorem 3.2, is that

$$|T \max_{\pi\in\mathcal{P}_K} \inf_{\lambda\in\mathbb{R}^d} \langle\pi, \mathcal{D}(\lambda)\rangle - \text{OPT}\,| \leqslant \mathcal{O}(\sqrt{T}),$$

therefore $\lim_{T\to\infty} \text{OPT}\,/T = \lim_{T\to\infty} \max_{\pi\in\mathcal{P}_K} \inf_{\lambda\in\mathbb{R}^d}\langle\pi, \mathcal{D}(\lambda)\rangle/T$ which also gives us a simple way to compute a close upper bound for OPT as a saddle point, instead of having to solve a combinatorial optimization problem.

Following the same arguments as in lemma 3.14 besides the last maximization step, we can also derive that

$$\text{static}-\text{OPT} \leqslant T \max_{k\in[K]} \inf_{\lambda\in\mathbb{R}^d} \mathcal{D}(\lambda, k).$$

It remains to find an example such that

$$\max_{k\in[K]} \inf_{\lambda\in\mathbb{R}^d} \mathcal{D}(\lambda, k) < \max_{\pi\in\mathcal{P}_K} \inf_{\lambda\in\mathbb{R}^d} \langle\pi, \mathcal{D}(\lambda)\rangle.$$

We used such an example in section 3.10.2, where $\max_{k\in[K]} \inf_{\lambda\in\mathbb{R}^d} \mathcal{D}(\lambda, k) = 0$ and $\max_{\pi\in\mathcal{P}_K} \inf_{\lambda\in\mathbb{R}^d}\langle\pi, \mathcal{D}(\lambda)\rangle = 0.25$. □

3.10.2 Simple example and numerical illustration

We illustrate proposition 3.1 and give an example that shows that randomizing over the different sources can outperform the best performance achievable by a single source.

We consider a case where the utility $u_t \in \{-1, 1\}$ and the protected attribute of a user $a_t \in \{-1, 1\}$. All combination of u_t and a_t are equiprobable: $\mathbb{P}(u_t = u, a_t = a) = 1/4$ for all $u, a \in \{-1, 1\}$. The penalty function is $R(\delta) = 5|\delta|$. There are $K = 2$ sources of information. The first source is capable of identifying good indivduals of group 1: it gives a context $c_{t1} = 1$ when $(a_t, u_t) = (1, 1)$ and $c_{t2} = 0$ otherwise. The second source does the same for good individuals of group -1: it gives $c_{t2} = 1$ when $(a_t, u_t) = (-1, 1)$ and $c_{t2} = 0$ otherwise. In terms of conditional expectations, this translates into $\mathbb{E}[u_t|c_{t1}{=}1] = \mathbb{E}[a_t|c_{t1}{=}1] = 1$, $\mathbb{E}[u_t|c_{t1}{=}0] = \mathbb{E}[a_t|c_{t1}{=}0] = -1/3$ for the first source; and $\mathbb{E}[u_t|c_{t2}{=}1] = 1$, $\mathbb{E}[a_t|c_{t2}{=}1] = -1$, $\mathbb{E}[u_t|c_{t2}{=}0] = -1/3$, $\mathbb{E}[a_t|c_{t2}{=}0] = 1/3$ for the second source.

We apply algorithm 1 to this example, and we compare its performance to (i) the one of an algorithm that has only access to the first source (which is the best source since both sources are symmetric), (ii) a greedy algorithm that only use the first source and selects $x_t = 1$ whenever $\mathbb{E}[u_t \mid c_{t1}] > 0$, and (iii) the offline optimal bounds OPT and static-OPT. The results are presented in fig. 3.1. We observe that our algorithm is close to OPT as shown in theorem 3.2. It also vastly outperforms the algorithm that simply picks the best source, whose performance is close to static-OPT. The greedy has a largely negative utility because it is unfair. Our randomizing algorithm picks an allocation $x_t = 1$ only when the context implies that $u_t = 1$. It obtain a fair allocation by switching sources. An algorithm that uses a single source cannot be fair unless it chooses $x_t = 0$, which is why static-OPT=0 for this example. The 1^{st} and 3^{rd} quartiles are reported for multiple random seeds, the solid lines are the means. The computations were done on a laptop with an i7-10510U and 16 Gb of ram.

3.10.3 Additional experiments on the effect on OPT of sources and fairness prices

In the main body of the paper, we study how one is able to use the online algorithm algorithm 1 so as to achieve a good total utility asymptotically close to that of the offline optimal OPT. However we only consider R and the prices p_k of the different sources as inputs for the algorithm, and we did not discuss how their variation actually affect the performance of OPT. While the previous illustrating example shows how a mixture of sources may achieve strictly better performance, we now only focus on the static optimization problem and how various metrics may change as we vary R and p_k.

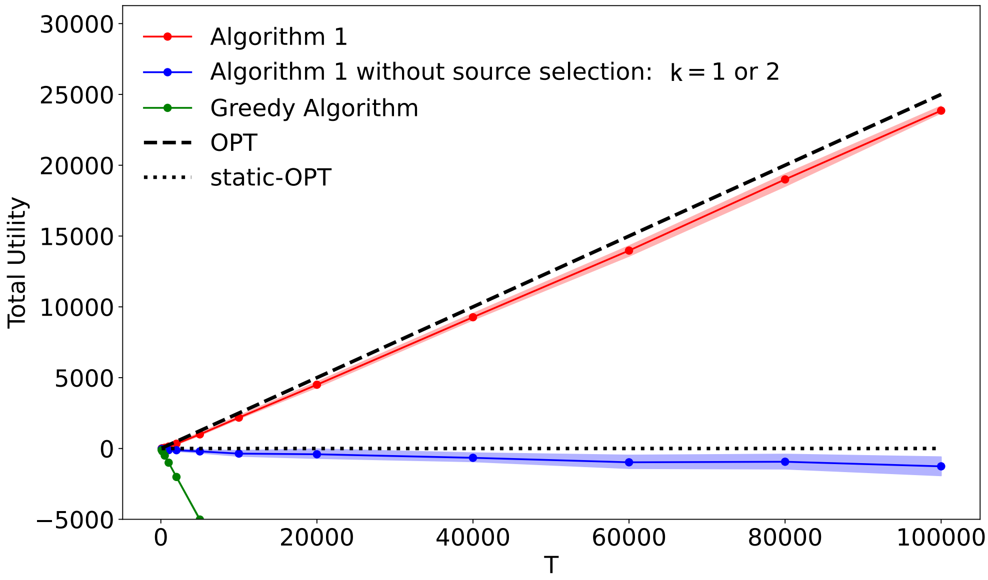


Fig. 3.1: Static source vs randomization: The red curve is the total utility of algorithm 1, The blue the utility of algorithm 1 ran on the best-fixed source, and the green the utility of an unfair greedy algorithm. Black curves correspond to the upper bounds $T \lim_{T\to\infty}(\text{OPT}/T)$ and similarly for static-OPT.

Data model: We consider a fairness penalty $R(x) = rx^2$ with $r \geqslant 0$, and $\mathcal{A} = \{-1, 1\}$ so that a_t can only be either -1 or 1, encoding men and women for instance. We suppose that the protected attributes and the utility (u_t, a_t) are distributed according to $\mathbb{P}(a_t = 1) = \mathbb{P}(a_t = -1) = 1/2$ and $u_t \sim \mathcal{N}(a_t, 1)$, hence the group for which $a_t = 1$ is generally correlated to having better utilities than those for which $a_t = -1$. We suppose that there are no public context, and that there are $K = 2$ information sources, with $c_{t,1} = (a_t, u_t)$ for source 1 associated with a price of $p_1 = p$, and for source $k = 2$ we can observe $c_{t,2} = u_t$ with a price of 0. Here the sources are 'monotone' in that one contains strictly more information than the other, but is more expensive. Basically we are going to observe the utility u_t in all cases, but before observing u_t we can pay p to additionally observe a_t. Note that when $r = 0$ then $R = 0$ and there is no fairness penalty meaning that we don't actually care about a_t and will always select source $k = 2$, while when $p = 0$ the information about a_t is free and there is no uncertainty.

We study the asymptotic regime with $T \to \infty$, and compute the optimal solution of the offline optimization problem through its dual problem. Figure 1 represent the quantities associated to the optimal solution for varying p and r: (top-left) π^* the percentage of times information is bought (i.e., $k = 1$ is chosen), (top-right) the probability of group 1 being selected $\mathbb{P}(a_t = 1 \mid x_t = 1)$, (bottom-left) the fairness cost relative to utility $R(\mathbb{E}[a_t x_t])/\mathbb{E}[u_t x_t]$, and (bottom-right) the information cost relative to utility $p\pi^*/\mathbb{E}[u_t x_t]$.

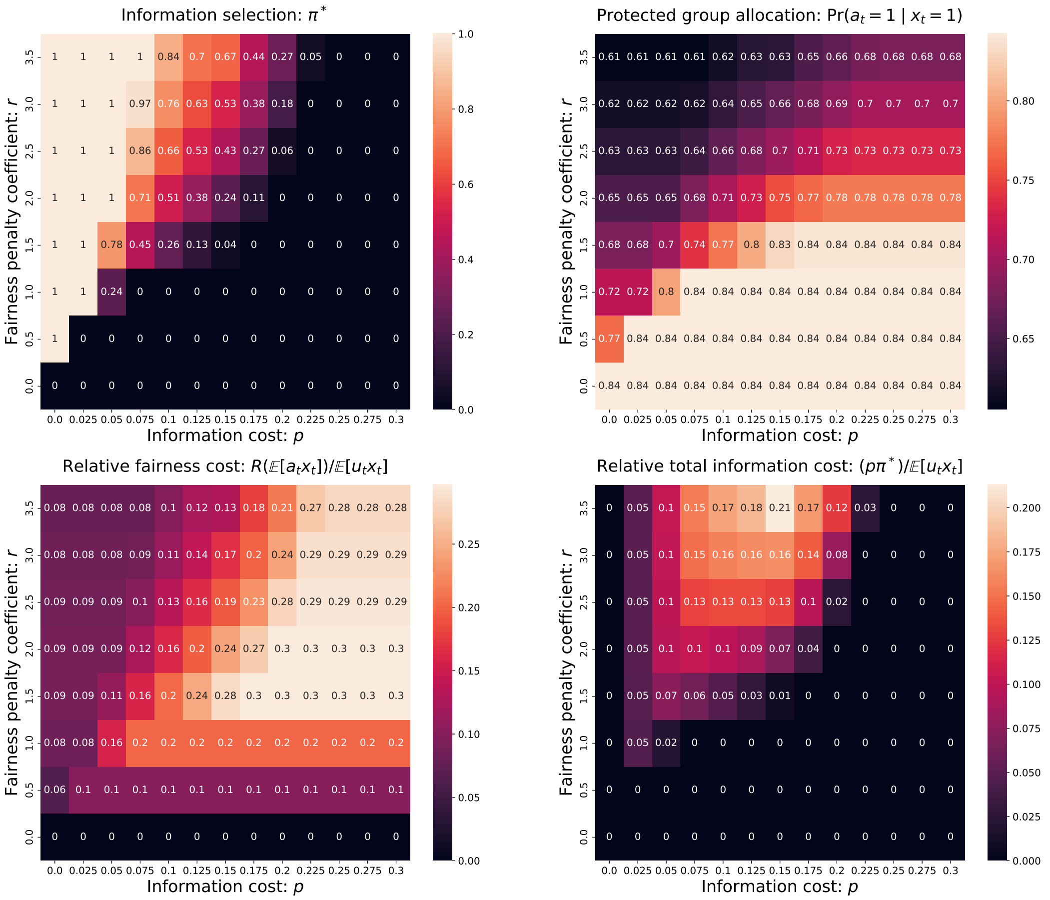


Fig. 3.2: For varying fairness scaling r and information cost p: (top-left) optimal information purchase frequency π^*, (top-right) optimal probability of group 1 selection $\mathbb{P}(a_t = 1 \,|\, x_t = 1)$, (bottom-left) optimal fairness cost relative to utility $R(\mathbb{E}[a_t x_t])/\mathbb{E}[u_t x_t]$, and (bottom-right) optimal information cost relative to utility $p\pi^*/\mathbb{E}[u_t x_t]$

Analysis and interpretation:

- *(top-left)* Looking at the optimal action in terms of information (to buy the observation of a_t or not), we see that there are three regimes: $\pi^* = 1$, $\pi^* = 0$ and a transitive regime $\pi^* \in (0, 1)$. When $r = 0$ there is no reason to care about a_t, thus we simply do not buy information ($\pi^* = 0$), this remains true as long as the cost of p is large over r. When $p = 0$ there are no costs to observe a_t and thus $\pi^* = 1$ and this remains true as long as the fairness penalty r is large enough over p. In between, we may obtain optimal actions which are strictly between 0 and 1: in this case the best action is a strict convex combination between the two sources and there is some trade-off between buying costly information and selecting an unfair allocation.

- *(top-right)* Regarding the proportion of individuals of group 1 selected, we see that it is always higher than 0.5 as group $a_t = 1$ correlates to higher utility compared to group $a_t = -1$. Therefore it is decreasing in r and goes towards the perfectly fair allocation 0.5 which would incur no penalty even for high r. It is also increasing in p as higher information costs make it less desirable to pay to observe a_t, and thus more difficult to accurately achieve a fair allocation: the support of the utility u_t conditioned on $a_t = 1$ and conditioned on $a_t = -1$ are not disjoint, which makes it difficult only by observing u_t to determine whether $a_t = 1$ or -1.

- *(bottom-left)* For the relative fairness cost compared to utility, when the information cost p increases the allocation becomes more unfair as it becomes more difficult without access to a_t to choose a fair allocation with good utility as describe above. Moreover the average utility u_t of the selected groups tend to decrease as the individuals which are more likely to be $a_t = -1$ based on u_t have low utility. Therefore the ratio fairness penalty over expected utility tend to increase. When r increases there are three conflicting effects: high r makes it more appealing to achieve a fair allocation at the cost of expected utility u_t which makes the latter decrease, the expected group allocation becomes more fair as seen previously, and the scaling of R increases. When r becomes high there is some balance between a fairer allocation and higher penalty, and when $r = 0$ the fairness penalty is clearly 0. Overall for each fixed p there is some maximum in terms of fairness penalty relative to utility.

- *(bottom-right)* Finally for the information cost relative to utility, it is increasing in r. Indeed due to large fairness penalties it is better to buy information thus π^* increases while p remains fixed, hence the product $p\pi^*$ increases. When p increases, we buy information less often (π^* decreases), and because $p\pi^* = 0$ for $p = 0$ or p large, there is some maximum for each r fixed.

Overall, we see that higher fairness penalties tend to make it more likely to buy information and select a fair allocation even at the cost of utility, while high information cost makes it unattractive to buy this information and can lead to unfairness due to the difficulty of identifying the protected attribute without this additional data.

Here we looked at a simple example to try to isolate how sensible is the optimal solution to some variations on r and p. Other parameters can lead to unfairness, such as an unbalance in the protected group population where one would still need a balanced allocation, or penalty functions with varying convexity strength.

3.11 Other Fairness Penalty — proof of theorem 3.7

In this section, we consider an other type of fairness penalty as described in section 3.4.1.

We redefine $\mathcal{U}(.)$ in the following way:

$$\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) = \sum_{t=1}^{T}(f_t(x_t) - p_{k_t}) - (\sum_{t=1}^{T} x_t)R\left(\frac{\sum_{t=1}^{T} a_t(x_t)}{\sum_{t=1}^{T} x_t}\right). \tag{3.30}$$

When $\mathcal{X} \subset \mathbb{R}_+^n$ with $n > 1$, there are two ways to intuitively generalize this penalty for higher dimension allocations. We can replace $\sum_{t=1}^{T} x_t$ by $\sum_{t=1}^{T}\sum_{i=1}^{n} x_{ti}$, which is still a scalar and the proof follows directly. For instance, this corresponds for online bipartite matching with node arrival to a fairness condition on all online arriving nodes (it does not matter to which offline nodes they are matched). Otherwise we can instantiate n penalties $R_1, \ldots, R_n$, using $\sum_t x_{ti}$ for each penalty R_i with $i \in [n]$. This corresponds to a fairness condition on the selected online nodes for each of the n offline node. This can be dealt with using one parameter λ_i for each penalty R_i. For this proof we will assume that $n = 1$ without loss of generality.

We require the following technical assumption: $\sum_{\tau=1}^{t} a_\tau(x_\tau)/\sum_{\tau=1}^{t} x_\tau \in \Delta$ for all x_τ, a_τ and t. Of course, the size of Δ can be increased to guarantee — if possible — that this quantity stays inside Δ, but with the trade-off that $\mathrm{Diam}(\Delta)$ increases. Clearly, this assumption is verified for the special case of $a_t(x_t) = ax_t$ for some $a \in \mathcal{A}$, as the $x_\tau/\sum_t x_t$ form a convex combination.

Let us redefine the virtual reward associated to this other fairness penalty:

$$\tilde{\varphi}(\lambda, c_{tk}, k) = \max_{x\in\mathcal{X}} \mathbb{E}\left([f_t(x) - \langle\lambda_t, a_t(x)\rangle \mid c_{tk}] + xR^*(\lambda_t)\right) - p_k.$$

We also redefine the unbiased estimator $\widehat{\varphi}$ with this new $\tilde{\varphi}$.

We now use line 0. There are 4 differences with line 0: the new function φ contains in addition $xR^*(\lambda)$, x_t is the $\arg\max$ with respect to this new virtual reward, γ_t is the $\arg\max$ over Δ_{δ_t/x_t} instead of Δ_{δ_t}, and finally the update for λ_{t+1} uses $x_t\gamma_t$ instead of γ_t.

Let us first prove an upper bound on OPT, which actually guides the design of the modified virtual reward $\tilde{\varphi}$:

Algorithm 3 Online Fair Allocation with Source Selection — Other fairness penalty

Input: Initial dual parameter λ_0, initial source-selection-distribution $\pi_0 = (1/K, \dots, 1/K)$, dual gradient descent step size η, EXP3 step size ρ, cumulative estimated rewards $S_0 = 0 \in \mathbb{R}^K$.

for $t \in [T]$ **do**

Draw a source $k_t \sim \pi_t$, where $(\pi_t)_k \propto \exp(\rho S_{(t-1)k})$ and observe c_{tk_t}.

Compute the allocation for user t:

$$x_t = \arg\max_{x \in \mathcal{X}} \left(\mathbb{E}[f_t(x) - \langle \lambda_t, a_t(x) \rangle \mid c_{tk_t}] + x R^*(\lambda_t) \right). \tag{3.31}$$

Update the estimated rewards sum and sources distributions for all $k \in [K]$:

$$S_{tk} = S_{t-1,k} + \widehat{\varphi}(\lambda_t, c_{tk}, k_t), \tag{3.32}$$

Compute the expected protected group allocation $\delta_t = \mathbb{E}[a_t(x_t) \mid c_{tk_t}, x_t]$ and compute the dual protected groups allocation target and update the dual parameter:

$$\gamma_t = \arg\max_{\gamma \in \Delta_{\delta_t / x_t}} \{ \langle \lambda_t, \gamma \rangle - \bar{R}(\gamma) \},$$
$$\lambda_{t+1} = \lambda_t - x_t \eta (\gamma_t - \delta_t / x_t).$$

end for

Lemma 3.16. *We have the following upper-bound for the offline optimum:*

$$\tilde{\mathrm{OPT}} \leqslant T \sup_{\pi \in \mathcal{P}_K} \inf_{\lambda \in \mathbb{R}^d} \langle \pi, \tilde{\mathcal{D}}(\lambda) \rangle,$$

where $\tilde{\mathcal{D}}(\lambda) = (\tilde{\mathcal{D}}(\lambda, 1), \dots, \tilde{\mathcal{D}}(\lambda, K))$ *is a vector representing the value of the dual conjugate problem, with for* $k \in [K]$ *the coordinates*

$$\tilde{\mathcal{D}}(\lambda, k) = \mathbb{E}_{c_{tk}}[\tilde{\varphi}(\lambda, c_{tk}, k)].$$

Proof. The idea will be the same as in lemma 3.14, we only use a different lagrangian function.

We recall that the x_t are $\sigma(c_{1k_1}, \dots, c_{tk_t})$ measurable (because the k_t are deterministic). Let $\delta_t = \mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk}]$, and $\tilde{f}_t(x) = \mathbb{E}_{f_t}[a_t(x_t) \mid c_{tk}]$. Using Jensen's inequality for R convex we have

$$\mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid c_{1k_1}, \dots, c_{Tk_T}]$$
$$= \sum_{t=1}^{T} \mathbb{E}[f_t(x_t) - p_{k_t} \mid c_{1k_1}, \dots, c_{Tk_T}] - \mathbb{E}[(\sum_{t=1}^{T} x_t) R \left(\frac{\sum_{t=1}^{T} a_t(x_t)}{\sum_{t=1}^{T} x_t} \right) \mid c_{1k_1}, \dots, c_{Tk_T}]$$

$$
\begin{aligned}
&= \sum_{t=1}^{T} \mathbb{E}[f_t(x_t) - p_{k_t} \mid c_{1k_1}, \dots, c_{Tk_T}] - (\sum_{t=1}^{T} x_t)\mathbb{E}[R\left(\frac{\sum_{t=1}^{T} a_t(x_t)}{\sum_{t=1}^{T} x_t}\right) \mid c_{1k_1}, \dots, c_{Tk_T}] \\
&\leqslant \sum_{t=1}^{T} \mathbb{E}[f_t(x_t) - p_{k_t} \mid c_{1k_1}, \dots, c_{Tk_T}] - (\sum_{t=1}^{T} x_t)R\left(\mathbb{E}[\frac{\sum_{t=1}^{T} a_t(x_t)}{\sum_{t=1}^{T} x_t} \mid c_{1k_1}, \dots, c_{Tk_T}]\right) \\
&= \sum_{t=1}^{T} \mathbb{E}[f_t(x_t) - p_{k_t} \mid c_{1k_1}, \dots, c_{Tk_T}] - (\sum_{t=1}^{T} x_t)R\left(\frac{\mathbb{E}[\sum_{t=1}^{T} a_t(x_t) \mid c_{1k_1}, \dots, c_{Tk_T}]}{\sum_{t=1}^{T} x_t}\right).
\end{aligned}
$$

Moreover by independence of the $(f_t, a_t, c_{t1}, \dots, c_{tK})$ we have that $\mathbb{E}[a_t(x_t) \mid \sigma(c_{1k_1}, \dots, c_{tk_t})] = \mathbb{E}_{a_t}[a_t(x_t) \mid c_{tk_t}] = \delta_t$. This is basically an application of lemma 3.11 where we state it more carefully. With the same argument we obtain that $\mathbb{E}[f_t(x_t) \mid \sigma(c_{1k_1}, \dots, c_{Tk_T})] = \tilde{f}_t(x_t)$. Therefore

$$
\mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid c_{1k_1}, \dots, c_{Tk_T}] \leqslant \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - (\sum_{t=1}^{T} x_t)R\left(\frac{\sum_{t=1}^{T} \delta_t}{\sum_{t=1}^{T} x_t}\right). \tag{3.33}
$$

We define the function $\mathcal{L} : (\boldsymbol{k}, \mathbf{c}, \mathbf{x}, \lambda) \mapsto \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - x_t\langle \lambda, \tilde{a}_t \rangle + x_t R^*(\lambda)$.

Using the Fenchel-Moreau theorem, and the previous inequality we have

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{k}, \mathbf{c}, \mathbf{x}, \lambda) &\geqslant \inf_{\lambda \in \mathbb{R}^d} \mathcal{L}(\mathbf{c}, \mathbf{x}, \lambda) \\
&= \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, \sum_{t=1}^{T} \tilde{a}_t x_t \rangle - (\sum_{t=1}^{T} x_t) R^*(\lambda)\} \\
&= \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - (\sum_{t=1}^{T} x_t) \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, \frac{\sum_{t=1}^{T} \delta_t}{\sum_{t=1}^{T} x_t} \rangle - R^*(\lambda)\} \\
&= \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - (\sum_{t=1}^{T} x_t) R^{**}\left(\frac{\sum_{t=1}^{T} \delta_t}{\sum_{t=1}^{T} x_t}\right) \\
&= \sum_{t=1}^{T} \tilde{f}_t(x_t) - p_{k_t} - (\sum_{t=1}^{T} x_t) R\left(\frac{\sum_{t=1}^{T} \delta_t}{\sum_{t=1}^{T} x_t}\right) \\
&\geqslant \mathbb{E}[\mathcal{U}(\boldsymbol{k}, \boldsymbol{x}) \mid \mathbf{c}_{\boldsymbol{k}}].
\end{aligned} \tag{3.34}
$$

We can then finish as we did for lemma 3.14. □

We now proceed to the rest of the proof of theorem 3.7:

Theorem. *For* $\eta = L/(2\,\mathrm{Diam}(\Delta)\sqrt{T})$, $m = \bar{f} + L + \max_k |p_k| + 2\eta\,\mathrm{Diam}(\Delta) + m^*$ *(*m^* *is a constant defined in the proof and depends on* Δ *and* R*),* $\rho = \sqrt{\log(K)/(TKm^2)}$*,* *and $\lambda_0 \in \partial R(a)$ for some $a \in \mathcal{A}$, algorithm algorithm 1 modified as suggested in section 3.4.1 has the following regret upper bound:*

$$\mathrm{Reg} \leqslant 2((L+\bar{f}+\max_k |p_k|+m^*)\sqrt{K\log(K)}+L\sqrt{d}+\mathrm{Diam}(\Delta))\sqrt{T}+2L\sqrt{K\log(K)}$$

Proof. We present here only the parts that are different from the proof of theorem 3.2, and we refer to the corresponding proof for more details.

Comparison of performance with $\tilde{\varphi}$ By adding and removing $-x_t\langle\lambda_t, \tilde{a}_t\rangle + x_t R^*(\lambda_t)$, and using that x_t is the maximizer that yields φ we derive that

$$\begin{aligned}\mathbb{E}[f_t(x_t)] &= \mathbb{E}[\mathbb{E}_{f_t}[f_t(x_t) \mid c_{tk_t}] - \langle\lambda_t, \delta_t\rangle + \langle\lambda_t, \delta_t\rangle + x_t R^*(\lambda_t) - x_t R^*(\lambda_t)] \\ &= \mathbb{E}[\varphi(\lambda_t, c_{tk_t}) + \langle\lambda_t, \delta_t\rangle - x_t R^*(\lambda_t)].\end{aligned}$$

Because $\Delta \subset \Delta_{\delta_t}$ and, because $\bar{R}$ and R are equal over Δ we have:

$$\begin{aligned}R^*(\lambda_t) &= \max_{\gamma\in\Delta}\{\langle\lambda_t, \gamma\rangle - R(\gamma)\} \\ &= \max_{\gamma\in\Delta}\{\langle\lambda_t, \gamma\rangle - \bar{R}(\gamma)\} \\ &\leqslant \max_{\gamma\in\Delta_{\tilde{a}_t}}\{\langle\lambda_t, \gamma\rangle - \bar{R}(\gamma)\} \\ &= \langle\lambda_t, \gamma_t\rangle - \bar{R}(\gamma_t).\end{aligned}$$

Thus

$$\mathbb{E}[f_t(x_t) - p_{k_t} - x_t\bar{R}(\gamma_t)] \geqslant \mathbb{E}[\tilde{\varphi}(\lambda_t, c_{tk_t}) + \langle\lambda_t, \delta_t - x_t\gamma_t\rangle]. \tag{3.35}$$

Application of bandit algorithms Compared to the previous setting, the only change is that the function R^* is included in φ, so we need to bound it as well. See lemma 3.13. This is where m^* comes from.

Dual gradient descent Let us now focus in the tracking of the fairness parameter. In the evaluation of our lower bound of the algorithm performance, we used $-x_t\bar{R}(\gamma_t)$ instead of the true penalty $(\sum_{t=1}^T x_t)R(\sum_{t=1}^T a_t(x_t)/\sum_{t=1}^T x_t)$. Because of a small modification of lemma 3.4 we only need to compare it with the penalty over the conditional expectation of $a_t x_t$. Indeed, by the Lipschitz property of R, whether we re-scale by T or $\sum x_t$ does not change anything.

Hence will now compare those two penalties. Here we will consider a modified version of $\bar{R}$, which is equal to $\bar{R}$ over $\bar{\Delta}$ and is equal to $+\infty$ outside of $\bar{\Delta}$, and we will use its convex conjugate $\bar{R}^*$ (note that the convex conjugate of $\bar{R}$ without the modification would have been different !). Let

$$\widehat{\lambda} \in \arg\max_{\lambda \in \mathbb{R}^d} \{ \langle \lambda, \frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \rangle - \bar{R}^*(\lambda) \}.$$

We have by Fenchel-Moreau theorem that

$$\begin{aligned} \langle \widehat{\lambda}, \frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \rangle - \bar{R}^*(\widehat{\lambda}) &= \bar{R}^{**} \left(\frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \right) \\ &= \bar{R} \left(\frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \right). \end{aligned}$$

Hence

$$\langle \widehat{\lambda}, \sum_{t=1}^T \delta_t \rangle - (\sum_{t=1}^T x_t) \bar{R} \left(\frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \right) = (\sum_{t=1}^T x_t) \bar{R}^*(\widehat{\lambda}).$$

By definition of the convex conjugate of $\bar{R}$, for any $\gamma \in \mathbb{R}^d$, thus in particular for all γ_t, we have

$$\bar{R}^*(\widehat{\lambda}) \geqslant \langle \widehat{\lambda}, \gamma_t \rangle - \bar{R}(\gamma_t).$$

Multiplying by x_t and summing for every $t \in [T]$ we obtain

$$(\sum_{t=1}^T x_t) \bar{R}^*(\widehat{\lambda}) \geqslant \sum_{t=1}^T \langle \widehat{\lambda}, x_t \gamma_t \rangle - x_t \bar{R}(\gamma_t)$$

Hence

$$\begin{aligned} & \langle \widehat{\lambda}, \sum_{t=1}^T \delta_t \rangle - (\sum_{t=1}^T x_t) \bar{R} \left(\frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \right) \geqslant \sum_{t=1}^T \langle \widehat{\lambda}, \gamma_t x_t \rangle - \bar{R}(\gamma_t) \\ \Leftrightarrow \quad & \langle \widehat{\lambda}, \sum_{t=1}^T \delta_t - x_t \gamma_t \rangle \geqslant (\sum_{t=1}^T x_t) \bar{R} \left(\frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \right) - \sum_{t=1}^T x_t \bar{R}(\gamma_t). \end{aligned}$$

Finally because $\bar{R}$ and R are equal over Δ, and because $\sum_{t=1}^T \delta_t / \sum_{t=1}^T x_t \in \Delta$ by the assumption, we have that

$$\langle \widehat{\lambda}, \sum_{t=1}^T \delta_t - x_t \gamma_t \rangle \geqslant (\sum_{t=1}^T x_t) R \left(\frac{\sum_{t=1}^T \delta_t}{\sum_{t=1}^T x_t} \right) - \sum_{t=1}^T x_t \bar{R}(\gamma_t) \tag{3.36}$$

As we did earlier, we apply Online Gradient Descent to track the term evaluated at $\widehat{\lambda}$ and obtain eq. (3.28). We just need to make sure that as before $\widehat{\lambda}$ is bounded. By the definition of $\widehat{\lambda}$:

$$\widehat{\lambda} \in \partial \bar{R}^{**}(\frac{\sum_{t=1}^{T} \delta_t}{\sum_{t=1}^{T} x_t}) = \partial \bar{R}(\frac{\sum_{t=1}^{T} \delta_t}{\sum_{t=1}^{T} x_t}) = \partial R(\frac{\sum_{t=1}^{T} \delta_t}{\sum_{t=1}^{T} x_t}).$$

Because R is L-Lipschitz for $\| \cdot \|_2$, we have $\|\widehat{\lambda}\|_2 \leqslant L$.

The rest of the proof is identical.

□

3.12 Learning u_t — proof of proposition 3.8

We prove in this section the result regarding the added regret of learning u_t. In terms of utility functions f_t, protected attributes a_t, and allocations x_t, we go back to the setting of section 3.2.1 with $f_t(x) = u_t x$ and $\mathcal{X} = \{0, 1\}$.

In addition to the assumptions made in section 3.4.2, we give more precision regarding the structure of u_t. We define $\theta_{tk} = \mathbb{1}[k_t = k]$ the indicator variable of selecting source k at time t. We define the following filtration for all $k \in [K]$ and all $t \in [T]$:

$$\mathcal{F}_{tk} = \sigma(x_1, c_{1k}, \theta_{1k}, u_1, \dots, x_t, c_{tk}, \theta_{tk}, u_t, x_{t+1}, c_{t+1,k}, \theta_{t+1,k}).$$

We assume that the contexts c_{tk} are now feature vectors of dimension q_k, and that for all $k \in [K]$, there exists some vector $\psi_k \in \mathbb{R}^{q_k}$ so that

$$\eta_{tk} = u_t - \langle \psi_k, c_{tk} \rangle, \tag{3.37}$$

is a zero mean 1-subgaussian random variable conditionally on $\mathcal{F}_{t-1,k}$.

Consider that each source k corresponds to one bandit problem, with an action set at time t composed of two contextual actions $\{c_{tk}\theta_{tk}x, x \in \mathcal{X}\} = \{c_{tk}\theta_{tk}, 0\}$, and a reward $u_t x_t \theta_{tk}$. If $\theta_{tk} = 0$ then the action selected is 0, and the new reward is 0, which means both are "observed". If $\theta_{tk} = 1$, then the feedback is u_t if $x_t = 1$, and 0 otherwise. Regardless of the source selected, the feedback received and the action producing this feedback are observed for all arms in parallel. This is akin to artificially playing all the $[K]$ bandits in simultaneously. This new reward still follows the conditional linearity assumption, and $x_t\theta_{tk}\eta_t$ is still $\mathcal{F}_{t-1,k}$ conditionally 1-subgaussian.

With these rewards and actions, we can define for each source $\widehat{\psi}_{tk}$ and V_{tk} the least squares estimator and the design matrix as done in Equations (19.5) and (19.6) of [LS20a]. Let $\delta \in (0,1)$. We define the confidence set parameter β_t as

$$\sqrt{\beta_t} = \bar{\psi} + \sqrt{2\log\left(\frac{1}{\delta}\right) + \max_{k\in[k]} q_k \log\left(1 + \frac{t\bar{c}^2}{\max_{k\in[k]} q_k}\right)}. \tag{3.38}$$

This parameter is independent of k, this will make the computations easier later on. We now define the ellipsoid confidence set of ψ_k as

$$\mathcal{I}_{tk} = \{\psi \in \mathbb{R}^{q_k}, \|\psi - \widehat{\psi}_{tk}\|^2_{V_{t-1,k}} \leqslant \beta_t\}. \tag{3.39}$$

algorithm 1 is modified in the following way. For each time t, let $\tilde{\psi}_{tk} = \arg\max_{\psi\in\mathcal{I}_{tk}} \langle \psi, c_{tk}\rangle$. We now select x_t using this optimistic estimate of ψ_k:

$$x_t = \arg\max_{x\in\{0,1\}} \{x(\langle\tilde{\psi}_{tk}, c_{tk}\rangle - \langle\lambda_t, \tilde{a}_t\rangle)\}. \tag{3.40}$$

Proposition. *Given the above assumptions, the* added *regret of having to learn* $\mathbb{E}[u_t \mid c_{tk}]$ *is of order* $\mathcal{O}(\sqrt{KT}\log(T))$.

The proof will proceed as follows: we first highlight where the learning will impact the algorithm, leverage the concentration results from [APS11] to decompose into a good event and bad event, and finally bound the loss over the good event.

Let us now analyze the main part of the proof in section 3.9 that is modified, which is "Comparison of performance with φ", because the only change is the quantity that x_t maximizes. Let x_t^* be the maximizer obtained by replacing the optimistic estimate with the true parameter ψ_k. Let $\tilde{a}_t = \mathbb{E}[a_t \mid c_{tk_t}]$. Because x_t maximizes the quantity with the optimistic estimate we have

$$\begin{aligned}
x_t\mathbb{E}[u_t \mid c_{tk_t}] - x_t\langle\lambda_t, \tilde{a}_t\rangle &= x_t\langle\psi_{k_t}, c_{tk_t}\rangle - x_t\langle\lambda_t, \tilde{a}_t\rangle \\
&= x_t(\langle\psi_{k_t}, c_{tk_t}\rangle - \langle\tilde{\psi}_{tk_t}, c_{tk_t}\rangle) + x_t\langle\tilde{\psi}_{tk_t}, c_{tk_t}\rangle - x_t\langle\lambda_t, \tilde{a}_t\rangle \\
&\geqslant x_t\langle\psi_{k_t} - \tilde{\psi}_{tk_t}, c_{tk_t}\rangle + x_t^*\langle\tilde{\psi}_{tk_t}, c_{tk_t}\rangle - x_t^*\langle\lambda_t, \tilde{a}_t\rangle \\
&= x_t\langle\psi_{k_t} - \tilde{\psi}_{tk_t}, c_{tk_t}\rangle + x_t^*\langle\tilde{\psi}_{tk_t} - \psi_{k_t}, c_{tk_t}\rangle + x_t^*\langle\psi_{k_t}, c_{tk_t}\rangle - x_t^*\langle\lambda_t, \tilde{a}_t\rangle \\
&= x_t\langle\psi_{k_t} - \tilde{\psi}_{tk_t}, c_{tk_t}\rangle + x_t^*\langle\tilde{\psi}_{tk_t} - \psi_{k_t}, c_{tk_t}\rangle + \varphi(\lambda_t, c_{tk_t}, k_t) + p_{k_t}.
\end{aligned}$$

The last term φ is the one we want, we now need to lower bound the first two terms.

By Theorem 20.5 of [LS20a], we know with probability at least $1-\delta$ that $\psi_k \in \mathcal{I}_{tk}$ for all t. By union bound, the probability that the ψ_k are simultaneously all in their confidence set is at least $1-K\delta$. We denote this event by I.

Suppose that the event I is satisfied. Then by definition of $\tilde{\psi}_{tk_t}$, we have

$$x_t^* \langle \tilde{\psi}_{tk_t} - \psi_{k_t}, c_{tk_t} \rangle \geqslant 0.$$

Now let us look at the last term. We can first decompose it over the K sources:

$$x_t \langle \psi_{k_t} - \tilde{\psi}_{tk_t}, c_{tk_t} \rangle = \sum_{k=1}^{K} \theta_{tk} x_t \langle \psi_k - \tilde{\psi}_{tk}, c_{tk} \rangle = \sum_{k=1}^{K} \langle \psi_k - \tilde{\psi}_{tk}, \theta_{tk} x_t c_{tk} \rangle.$$

Now, consider one source k, and sum the difference term for every t. By remarking that $\theta_{tk} = \theta_{tk}^2$, we have that

$$\begin{aligned}
\sum_{t=1}^{T} |\langle \psi_k - \tilde{\psi}_{tk} \mid \theta_{tk} x_t c_{tk} \rangle| &= \sum_{t=1}^{T} \theta_{tk} |\langle \psi_k - \tilde{\psi}_{tk} \mid \theta_{tk} x_t c_{tk} \rangle| \\
&\leqslant \sum_{t=1}^{T} \theta_{tk} \|\psi_k - \tilde{\psi}_{tk}\|_{V_{t-1,k}} \|\theta_{tk} x_t c_{tk}\|_{V_{t-1,k}^{-1}} \\
&\qquad \text{(dual-norm inequality)} \\
&\leqslant 2 \sum_{t=1}^{T} \theta_{tk} \sqrt{\beta_t} \|\theta_{tk} x_t c_{tk}\|_{V_{t-1,k}^{-1}} \qquad \text{(by the event } I) \\
&\leqslant 2 \sum_{t=1}^{T} \theta_{tk} \sqrt{\beta_T} \max\{1, \|\theta_{tk} x_t c_{tk}\|_{V_{t-1,k}^{-1}}\} \\
&\qquad (\beta_t \text{ are increasing}) \\
&\leqslant 2\sqrt{\beta_T} \sqrt{\sum_{t=1}^{T} \theta_{tk}^2} \sqrt{\sum_{t=1}^{T} \max\{1, \|\theta_{tk} x_t c_{tk}\|_{V_{t-1,k}^{-1}}^2\}} \\
&\qquad \text{(C.S. inequality)} \\
&= 2\sqrt{\beta_T} \sqrt{\sum_{t=1}^{T} \theta_{tk}} \sqrt{\sum_{t=1}^{T} \max\{1, \|\theta_{tk} x_t c_{tk}\|_{V_{t-1,k}^{-1}}^2\}} \\
&= 2\sqrt{\beta_T} \sqrt{T_k} \sqrt{\sum_{t=1}^{T} \max\{1, \|\theta_{tk} x_t c_{tk}\|_{V_{t-1,k}^{-1}}^2\}},
\end{aligned}$$

where T_k is the number of times source k is selected.

By Lemma 19.4 of [LS20a], we have that

$$\sum_{t=1}^{T} \max\{1, \|\theta_{tk} x_t c_{tk}\|^2_{V^{-1}_{t-1,k}}\} \leqslant 2q_k \log\left(1 + \frac{T\bar{c}^2}{q_k}\right).$$

Therefore, using Jensen inequality for the square root function which is concave over the T_k, and because $\sum_k T_k = T$, we obtain

$$\begin{aligned}
\sum_{t=1}^{T}\sum_{k=1}^{K} |\langle \psi_k - \tilde{\psi}_{tk} \mid \theta_{tk} x_t c_{tk}\rangle| &\leqslant \sum_{k=1}^{K} 2\sqrt{\beta_T}\sqrt{T_k}\sqrt{2q_k \log\left(1 + \frac{T\bar{c}^2}{q_k}\right)} \\
&\leqslant 2\sqrt{\beta_T}\sum_{k=1}^{K} \sqrt{T_k}\sqrt{2q_k \log\left(1 + \frac{T\bar{c}^2}{q_k}\right)} \\
&\leqslant \sqrt{8T\beta_T}\sqrt{\sum_{k=1}^{K} q_k \log\left(1 + \frac{T\bar{c}^2}{q_k}\right)}.
\end{aligned}$$

This yields a $\mathcal{O}(\sqrt{KT}\log(T))$ error.

And with probability at most $K\delta$ the concentration event I does not hold, but the error is bounded by 4 for each t. Thus overall in expectation, it is bounded by $4TK\delta$. Picking $\delta = 1/T$, yields the desired sub-linear regret bound.

3.13 Learning a_t

In this section, we provide further intuitions on the existence of $[K]$ different sources, and how it is possible to adapt the algorithm to learn $\mathbb{E}[a_t \mid c_{tk}]$ on the fly. Let us consider the special case of our model, where the additional information given by each source is a noisy observation $\widehat{a}_{tk}$ of a_t, with different noise levels; in the main setting this can be described as $c_{tk} = (z_t, \widehat{a}_{tk})$. Sources with lower noise levels will have higher prices (one directly pays for the level of accuracy). This setting can be used to give users the level of privacy they are willing to give up (as in Local Differential Privacy [DR14]), by revealing some noisy information about their private data; of course, the more data is revealed (or the smaller the noise), the higher the compensation (the price).

Furthermore, if the source k corresponds to adding additive independent noise L_{tk} and to returning $\widehat{a}_{tk} = a_t + L_{tk}$, then knowing the law of both L_{tk} and a_t, the decision maker can directly compute the conditional expectation of a_t having observed $\widehat{a}_{tk}$. Nevertheless, if she only knows the noise distribution, but not the law of a_t, conditional expectation can still be learned in an online fashion in some cases.

Indeed, suppose for simplicity that $a_t \in \{-1, 1\}$, $\mathbb{P}(a_t = 1) = \alpha \in (0, 1)$, (z_t, u_t) is independent of a_t (therefore we can focus only in dealing with a_t), and that the L_{tk} are 0 mean σ_k sub-exponential random variables (see 2.7.5 [Ver18]). We define the following running empirical estimate of α:

$$\widehat{\alpha}_t = \frac{1}{t}\sum_{\tau=1}^{t}(\widehat{a}_{\tau,k_\tau} + 1)/2.$$

It is an unbiased estimator of α as

$$\mathbb{E}[\widehat{a}_{tk}] = \mathbb{E}[a_t] + \mathbb{E}[L_{tk}] = 2\alpha - 1 + 0 = 2\alpha - 1.$$

Note that the $\widehat{a}_{t,k_t}$ are *not* independent, and the conditional expectation is *not* Lispchitz-continuous with respect to the parameter α. Still, we can derive an algorithm with a sub-linear regret bound.

First let us state a Lemma on the concentration of $\widehat{\alpha}_t$

Lemma 3.17. *We have the following concentration bound for $t \gtrsim \log(T)$:*

$$\mathbb{P}\left(|\widehat{\alpha}_t - \alpha| \geqslant \sqrt{\frac{\kappa^2\log(T)}{ct}}\right) \leqslant \frac{2}{T} \tag{3.41}$$

Proof. We refer to Chapter 2 of [Ver18] for concentration results of sum of independent random variables. Let us recall some definitions, properties, and a concentration inequality.

Let us consider the following Orlicz space norm for random variables. For a random variable X, we define

$$\|X\|_{\psi^1} = \inf\{t > 0, \mathbb{E}[\exp(|X|/t)] \leqslant 2\}.$$

If $\|X\|_{\psi^1}$ is finite, then X is called sub-exponential. It is a norm, and for a bounded variable X in $[-1, 1]$, we have $\|X\|_{\psi^1} \leqslant 1/\log(2)$ (see exercise 2.5.7, example 2.5.8, and Lemma 2.7.6 of [Ver18]). Hence because $(a_t + 1)/2 - \alpha \in [-\alpha, 1-\alpha] \subset [-1, 1]$, for $\widehat{\alpha}_{tk}$ the estimate obtained by always choosing source k, we have by triangle inequality that

$$\|\widehat{\alpha}_{tk} - \alpha\|_{\psi^1} \leqslant \frac{1}{\log(2)} + \frac{\sigma_k}{2}.$$

We now recall Bernstein's inequality:

Theorem (Theorem 2.8.1 of [Ver18]). *Let $X_1, \ldots, X_N$ be N independent, mean 0, sub-exponential random variables, with $\kappa = \max_i \|X_i\|_{\psi^1}$. Then for $c > 0$ a constant, for every $t > 0$ we have*

$$\mathbb{P}\left\{\left|\sum_{i=1}^{N} X_i\right| \geqslant t\right\} \leqslant 2\exp\left[-c\min\left(\frac{t^2}{N\kappa^2}, \frac{t}{\kappa}\right)\right].$$

We want to apply this Theorem to $(\widehat{a}_{t,k_t}+1)/2-\alpha$ the centered $\widehat{\alpha}_t$ where we now change the source depending on k_t. Remark that now however the $\widehat{\alpha}_t$ is not a sum of independent variables anymore because of k_t which does depend on the previous realizations of the $\widehat{a}_{t,k_t}$. Nevertheless it is a martingale difference for the filtration $\mathcal{H}_t$, and similar results will apply. We denote by $\kappa = 1/\log(2) + \max_k \sigma_k/2$ an upper bound on the sub-exponential norm of the centered $\widehat{\alpha}_t$ conditional on $\mathcal{H}_{t-1}$. Then the martingale $\sum_{\tau=1}^{t} \widehat{\alpha}_\tau - \alpha$ is also a sub-exponential random variable.

This is a known result, which is stated here for completeness, but that can otherwise be skipped. Indeed from proposition $2.7.1$ part $e)$ of [Ver18], we know that finite sub-exponential norm for a centered random variable X means that the moment generating function of X at y is bounded by $\exp(c_1\|X\|_{\psi^1}^2 y^2)$ for $y \leqslant c_2/\|X\|_{\psi^1}$, with c_1 and c_2 some positive constants.

Therefore, for $y \leqslant c_2/\kappa$, using the property of the conditional sub-exponentiality of $\widehat{\alpha}_t$, we have

$$\begin{aligned}\mathbb{E}[\exp(y\sum_{\tau=1}^{t}(\widehat{\alpha}_\tau-\alpha))] &= \mathbb{E}[\mathbb{E}[\exp(y\sum_{\tau=1}^{t}(\widehat{\alpha}_\tau-\alpha)) \mid \mathcal{F}_{t-1}]] \\ &= \mathbb{E}[\exp(y\sum_{\tau=1}^{t-1}(\widehat{\alpha}_\tau-\alpha))\mathbb{E}[\exp(y(\widehat{\alpha}_t-\alpha)) \mid \mathcal{F}_{t-1}]] \\ &\leqslant \mathbb{E}[\exp(y\sum_{\tau=1}^{t-1}(\widehat{\alpha}_\tau-\alpha))\mathbb{E}[\exp(c_1 y^2\kappa^2) \mid \mathcal{F}_{t-1}]].\end{aligned}$$

Iterating this inequality we obtain for $y \leqslant c_2/\kappa$ that

$$\mathbb{E}[\exp(y\sum_{\tau=1}^{t}(\widehat{\alpha}_\tau-\alpha))] \leqslant \mathbb{E}[\exp(c_1 t\kappa^2 y^2)].$$

This yields inequality (2.24) in [Ver18] the main element needed for the proof of Bernstein's Inequality. Thus Bernstein's inequality still holds for centered martingale sequence which are conditionally sub-exponential. We stress that this is not a new concentration inequality result.

Let us now apply Bernstein's Inequality with a deviation of order $\sqrt{\kappa^2 \log(T)/(ct)}$, for $t \geqslant \log(T)/c$:

$$\begin{aligned}
\mathbb{P}\left(|\widehat{\alpha}_t - \alpha| \geqslant \sqrt{\frac{\kappa^2 \log(T)}{ct}}\right) &= \mathbb{P}\left(|\sum_{\tau=1}^{t}(\widehat{\alpha}_\tau - \alpha)| \geqslant \sqrt{\frac{\kappa^2 t \log(T)}{c}}\right) \\
&\leqslant 2\exp\left[-c\min\left(\frac{\log(T)}{c}, \sqrt{\frac{\log(T)t}{c}}\right)\right] \\
&= \frac{2}{T}
\end{aligned}$$

□

Proposition 3.18. *Given these assumptions, using algorithm 1 with $\widehat{\alpha}_t$ instead of α to compute the conditional expectations yields an added regret of order $\mathcal{O}((\max_k \sigma_k + 2/\log(2))\sqrt{T\log(T)})$.*

Proof. There are two parts of the proof of theorem 3.2 that we need to modify, when the x_t is involved, and when we compare the penalty. We will deal with the first part (the more complicated), and the second one follows immediately by the same arguments. This will be done in 3 steps: we first highlight where the learning error occurs, then use the concentration result to decompose the error under a good and bad event, and finally show that under the good event the error in expectation is small enough.

Impact of learning error We suppose for simplicity that L_{tk} has a density g_k over the support $\mathbb{R}$, but this also holds for discrete random variables. We denote by ν_k the density of $\widehat{a}_{tk}$. First let us express $\tilde{a}_t = \mathbb{E}[a_t \mid c_{tk_t}]$ as a function of $\widehat{a}_{t,k_t} = \widehat{a}$ and α:

$$\begin{aligned}
\mathbb{E}[a_t \mid \widehat{a}_{t,k_t} = \widehat{a}] &= \mathbb{P}(a_t = 1 \mid \widehat{a}_{t,k_t} = \widehat{a}) - \mathbb{P}(a_t = -1 \mid \widehat{a}_{t,k_t} = \widehat{a}) \\
&= \frac{\mathbb{P}(a_t = 1)g_{k_t}(\widehat{a} - 1)}{\nu_{k_t}(\widehat{a})} - \frac{\mathbb{P}(a_t = -1)g_{k_t}(\widehat{a} + 1)}{\nu_{k_t}(\widehat{a})} \\
&= \frac{\alpha g_{k_t}(\widehat{a} - 1) - (1-\alpha)g_{k_t}(\widehat{a} + 1)}{\alpha g_{k_t}(\widehat{a} - 1) + (1-\alpha)g_{k_t}(\widehat{a} + 1)}.
\end{aligned}$$

Let us define the plug-in estimate of $\tilde{a}_t$ using $\bar{\alpha}_{t-1}$ whenever $\bar{\alpha}_{t-1}$ is in $[0,1]$ by:

$$s_t(\widehat{a}) = \frac{\bar{\alpha}_{t-1} g_{k_t}(\widehat{a}-1) - (1-\bar{\alpha}_{t-1}) g_{k_t}(\widehat{a}+1)}{\bar{\alpha}_{t-1} g_{k_t}(\widehat{a}-1) + (1-\bar{\alpha}_{t-1}) g_{k_t}(\widehat{a}+1)}. \tag{3.42}$$

When the estimated parameter is not in $[0,1]$, we use $s_t = 0$.

We define by x_t the maximisation of the virtual value function φ when replacing $\tilde{a}_t$ by s_t, and by x_t^* the allocation obtained if we were to actually use $\tilde{a}_t$. Because x_t is a maximizer of $u_t x_t - \langle \lambda_t, s_t \rangle x_t$, we have

$$\begin{aligned} u_t x_t &= u_t x_t - \langle \lambda_t, s_t \rangle x_t + \langle \lambda_t, s_t \rangle x_t \\ &\geqslant u_t x_t^* - \langle \lambda_t, s_t \rangle x_t^* + \langle \lambda_t, s_t \rangle x_t \\ &= \varphi(\lambda_t, \widehat{a}_{t,k_t}, k_t) + \langle \lambda_t, \tilde{a}_t - s_t \rangle x_t^* + \langle \lambda_t, s_t \rangle x_t. \end{aligned}$$

The term $\langle \lambda_t, s_t \rangle x_t$ will be tracked with the help of the OGD, it remains to take care of $\langle \lambda_t, \tilde{a}_t - s_t \rangle x_t^*$. By Cauchy Schwartz:

$$|\langle \lambda_t, \tilde{a}_t - s_t \rangle x_t^*| \leqslant \|\lambda_t\|_2 \|\tilde{a}_t - s_t\|_2.$$

We know that λ_t is already bounded, so we only need to bound the right-hand term. Decomposing depending on the values of k_t, we can rewrite

$$s_t - \tilde{a}_t = \sum_{k=1}^{K} \mathbb{1}[k_t = k](s_t - \mathbb{E}[a_t \mid \widehat{a}_{tk}])$$

Decomposition under good and bad event Denote by $\mathcal{C}_t = [|\widehat{\alpha}_t - \alpha| \leqslant \sqrt{(\kappa^2 \log(T)/(ct)}]$ the event of $\widehat{\alpha}_t$ concentrating around its mean. For t large enough, that is to say $t \geqslant (\kappa^2 \log(T))/(c \min(\alpha, 1-\alpha))$, the event $\mathcal{C}_t$ is included in $[\bar{\alpha}_{t-1} \in (0,1)]$. We can then decompose the analysis over these events using the triangle inequality: $|s_t - \tilde{a}_t| \leqslant \mathbb{1}_{\mathcal{C}_{t-1}} |s_t - \tilde{a}_t| + \mathbb{1}_{\bar{\mathcal{C}}_{t-1}} |s_t - \tilde{a}_t|$, and we know that for the term multiplied by $\mathbb{1}_{\mathcal{C}_{t-1}}$ we have that either it is 0 or $\bar{\alpha}_{t-1} \in [0,1]$. Let us consider such a $t-1$ big enough, and first analyze the left-hand term with the "good" event.

Small error under good event in expectation We consider the conditional expectation of this sum given k_t, and $\bar{\alpha}_{t-1}$. By independence between $(k_t, \bar{\alpha}_{t-1}, E)$ and $\widehat{a}_{tk}$ for all k, we have:

$$\mathbb{E}[\mathbb{1}_{\mathcal{C}_{t-1}}|s_t - \tilde{a}_t| \mid k_t, \bar{\alpha}_{t-1}]$$
$$\leqslant \mathbb{1}_{\mathcal{C}_{t-1}} \sum_{k=1}^{K} \mathbb{1}[k_t = k] \int_{\mathbb{R}} \left| \frac{\bar{\alpha}_{t-1} g_k(\widehat{a}-1) - (1-\bar{\alpha}_{t-1}) g_k(\widehat{a}+1)}{\bar{\alpha}_{t-1} g_k(\widehat{a}-1) + (1-\bar{\alpha}_{t-1}) g_k(\widehat{a}+1)} - \frac{\alpha g_k(\widehat{a}-1) - (1-\alpha) g_k(\widehat{a}+1)}{\alpha g_k(\widehat{a}-1) + (1-\alpha) g_k(\widehat{a}+1)} \right| \nu_k(\widehat{a}) d\widehat{a}.$$

Because $\nu_{k_t}(\widehat{a}) = (\alpha g_k(\widehat{a}-1) + (1-\alpha) g_k(\widehat{a}+1))$ we can simplify the integral:

$$\mathbb{E}[\mathbb{1}_{\mathcal{C}_{t-1}}|s_t - \tilde{a}_t| \mid k_t, \bar{\alpha}_{t-1}] \leqslant \mathbb{1}_{\mathcal{C}_{t-1}} \sum_{k=1}^{K} \mathbb{1}[k_t = k] \int_{\mathbb{R}} 2|\bar{\alpha}_{t-1} - \alpha| \frac{g_k(\widehat{a}-1) g_k(\widehat{a}+1)}{\bar{\alpha}_{t-1} g_k(\widehat{a}-1) + (1-\bar{\alpha}_{t-1}) g_k(\widehat{a}+1)} d\widehat{a}.$$

For $\alpha \in [0, 1]$, we have the following inequality:

$$\frac{xy}{\alpha x + (1-\alpha) y} \leqslant \max(x, y) \leqslant x + y.$$

Thus using this inequality and the fact that g_k is a density (which sums to 1), we obtain that

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_{\mathcal{C}_{t-1}}|s_t - \tilde{a}_t| \mid k_t, \bar{\alpha}_{t-1}] &\leqslant 2\mathbb{1}_{\mathcal{C}_{t-1}} |\bar{\alpha}_{t-1} - \alpha| \sum_{k=1}^{K} \mathbb{1}[k_t = k] \Big(\int_{\mathbb{R}} g_k(\widehat{a}+1) d\widehat{a} + \int_{\mathbb{R}} g_k(\widehat{a}-1) d\widehat{a} \Big) \\
&\leqslant 4\mathbb{1}_{\mathcal{C}_{t-1}} |\bar{\alpha}_{t-1} - \alpha| \sum_{k=1}^{K} \mathbb{1}[k_t = k] \\
&\leqslant 4\sqrt{\frac{\kappa^2 \log(T)}{c(t-1)}} \sum_{k=1}^{K} \mathbb{1}[k_t = k].
\end{aligned}$$

Finally taking the full expectation yields

$$\mathbb{E}[\mathbb{1}_{\mathcal{C}_{t-1}}|s_t - \tilde{a}_t|] \leqslant 4\sqrt{\frac{\kappa^2 \log(T)}{c(t-1)}}.$$

For the complementary event $\bar{\mathcal{C}}_{t-1}$:

$$\mathbb{E}[\mathbb{1}_{\bar{\mathcal{C}}_{t-1}}|s_t - \tilde{a}_t|] \leqslant 2\mathbb{E}[\mathbb{1}_{\bar{\mathcal{C}}_{t-1}}] = \frac{4}{T}.$$

Putting everything back together Now denoting by m_λ the upper bound over the λ_t obtained by lemma 3.3 and summing over all t:

$$
\begin{aligned}
|\mathbb{E}[\sum_{t=1}^{T}\langle\lambda_t, \tilde{a}_t - s_t\rangle x_t^*]| &\leqslant m_\lambda \sum_{t=1}^{T} \mathbb{E}[|s_t - \tilde{a}_t|] \\
&\leqslant 2m_\lambda \frac{\kappa^2 \log(T)}{c\min(\alpha, 1-\alpha)} + m_\lambda \sum_{t=\frac{\kappa^2\log(T)}{c\min(\alpha,1-\alpha)}+1}^{T} \mathbb{E}[|s_t - \tilde{a}_t|] \\
&\leqslant 2m_\lambda \frac{\kappa^2 \log(T)}{c\min(\alpha, 1-\alpha)} + m_\lambda \sum_{t=\frac{\kappa^2\log(T)}{c\min(\alpha,1-\alpha)}+1}^{T} (4\sqrt{\frac{\kappa^2\log(T)}{c(t-1)}} + \frac{4}{T}) \\
&\leqslant 2m_\lambda \frac{\kappa^2 \log(T)}{c\min(\alpha, 1-\alpha)} + m_\lambda \sum_{t=1}^{T} (4\sqrt{\frac{\kappa^2\log(T)}{ct}} + \frac{4}{T}) \\
&\leqslant m_\lambda (\frac{8}{\sqrt{c}}\sqrt{\kappa^2 T\log(T)} + 2\frac{\kappa^2\log(T)}{c\min(\alpha,1-\alpha)} + 4) = \mathcal{O}(\sqrt{T\log(T)}),
\end{aligned}
$$

where we used a series-integral comparison of $x \mapsto 1/\sqrt{x}$ for the second to last inequality.

There is a second place where replacing $\tilde{a}_t$ by s_t impacts the performance of the algorithm. Indeed, through the gradient descent on λ_t, we will have a penalty converging towards $R(\sum_{t=1}^{T} s_t x_t/T)$ and not $R(\sum_{t=1}^{T} \tilde{a}_t x_t/T)$. This is because the Online Gradient Descent for the λ_t uses s_t and not $\tilde{a}_t$. Nevertheless, because of the Lipschitz property of R:

$$
\mathbb{E}[T|R(\frac{\sum_{t=1}^{T}\tilde{a}_t x_t}{T}) - R(\frac{\sum_{t=1}^{T} s_t x_t}{T})|] \leqslant L\sum_{t=1}^{T}\mathbb{E}[|\tilde{a}_t - s_t|] \leqslant \mathcal{O}(\sqrt{T\log(T)}),
$$

where the last inequality is obtained by similar manipulations on $\mathcal{C}_t$ to what we just did.

Overall, the added regret when using s_t instead of $\tilde{a}_t$ is of order $\sqrt{T\log(T)}$ which still guarantees a sub-linear regret. □

Another algorithm is possible: first sample the source with the lowest sub-exponential constant $T^{2/3}\log(T)$ times, and then use the estimate obtained to compute s_t (without updating it with the newest noise received). The advantage is that even though we incur a $T^{2/3}$ regret (this can be shown using basically the same methods), we can obtain a lower κ. Hence if there is a big difference between the sub-exponential constant of the different sources, or even one source with infinite sub-exponential constant, this may be a viable algorithm.

3.14 Public Contexts

For this section and in order to make the dependence in the public context z_t appear more clearly, we suppose that c_{tk} represent only the additional information of the source k, and thus (c_{tk_t}, z_t) represent the whole information available at time t after selecting a source.

3.14.1 Finite number of Contexts — proof of proposition 3.5

We propose the following line 0.

Algorithm 4 Online Fair Allocation with Source Selection — With public contexts

---

Input: Initial dual parameter λ_0, initial source-selection-distribution $\pi_0(z) = (1/K, \dots, 1/K)$, dual gradient descent step size η, EXP3 step size $\rho_t(z)$ that uses a doubling trick, cumulative estimated rewards $S_0(z) = 0 \in \mathbb{R}^K$ for all $z \in \mathcal{Z}$.

for $t \in [T]$ **do**

Observe $z_t \in \mathcal{Z}$

Draw a source $k_t \sim \pi_t(z_t)$, where $\pi_{tk}(z_t) \propto \exp(\rho_t(z_t) S_{(t-1)k}(z_t))$ and observe c_{tk_t}.

Compute the allocation for user t:

$$x_t = \arg\max_{x \in \mathcal{X}} \mathbb{E}[f_t(x) - \langle \lambda_t, a_t(x) \rangle \mid c_{tk_t}, z_t].$$

Update the estimated rewards sum and sources distributions for all $k \in [K]$:

$$S_{tk}(z_t) = S_{t-1,k}(z_t) + \widehat{\varphi}(\lambda_t, c_{tk}, k_t, z_t),$$
$$S_{tk}(z) = S_{t-1,k}(z), \quad \forall z \in \mathcal{Z} \setminus \{z_t\}.$$

Compute the expected protected group allocation $\delta_t = \mathbb{E}[a_t(x_t) \mid c_{tk_t}, z_t, x_t]$ and compute the dual protected groups allocation target and update the dual parameter:

$$\gamma_t = \arg\max_{\gamma \in \Delta_{\delta_t}} \{\langle \lambda_t, \gamma \rangle - \bar{R}(\gamma)\},$$
$$\lambda_{t+1} = \lambda_t - \eta(\gamma_t - \delta_t).$$

end for

Proposition. *For μ the probability distribution over $\mathcal{Z}$ finite, we can derive an algorithm that has a modified regret of order $\mathcal{O}(\sqrt{TK\log(K)}\sum_{z\in\mathcal{Z}}\sqrt{\mu(z)})$, where $\mu(z)$ is the probability that the public attribute is z.*

Proof. Most of the proof consists is identical to the proof of theorem 3.2, and only the different parts will be highlighted, which is the upper bound of OPT, and applying bandits algorithms.

We first upper bound the performance of OPT. For all t let $h_t \in [K]^{\mathcal{Z}}$ be the deterministic source selection policies, $k_t = h_t(z_t)$ the selected sources, $x_t \in \mathcal{X}$ be (z_t, c_{t,k_t}) measurable, and $\lambda \in \mathbb{R}^d$ be a dual variable. In an identical manner to eq. (3.20) we can obtain the following inequality:

$$\max_{\mathbf{x}\in\mathcal{X}^T} \mathbb{E}[\mathcal{U}(\boldsymbol{x},\boldsymbol{k}) \mid z_1, c_{1,k_1}, \dots, z_T, c_{T,k_T}] \leqslant \sum_{t=1}^{T} \max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[f_t(x) - \langle \lambda, a_t(x)\rangle \mid z_t, c_{tk_t}] - p_{k_t} + TR^*(\lambda).$$

Let us define $\varphi(\lambda, c_{tk}, k, z) = \max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}[f_t(x) - \langle \lambda, a_t(x)\rangle \mid z, c_{tk}] - p_k$, and $\mu(z) = \mathbb{P}(z_t = z)$. We also define $\pi_k(z) = \sum_{t=1}^{T} \mathbb{1}[h_t(z) = k]/T$ the total number of times we would have selected source k for the public context z.

We now take the expectation of the sum. We obtain by tower property of the conditional expectation, and using the i.i.d. assumption:

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}[\varphi(\lambda, c_{t,h_t(z_t)}, h_t(z_t), z_t)] &= \sum_{t=1}^{T} \mathbb{E}[\mathbb{E}[\varphi(\lambda, c, h_t(z_t), z_t) \mid z_t]] \\
&= \sum_{t=1}^{T} \sum_{z\in\mathcal{Z}} \mathbb{E}[\varphi(\lambda, c, h_t(z), z) \mid z]\mu(z) \\
&= \sum_{z\in\mathcal{Z}} \sum_{t=1}^{T} \mathbb{E}[\varphi(\lambda, c, h_t(z), z) \mid z]\mu(z) \\
&= \sum_{z\in\mathcal{Z}} \sum_{k\in[K]} \sum_{\substack{t=1 \\ h_t(z)=k}}^{T} \mathbb{E}[\varphi(\lambda, c, k, z) \mid z]\mu(z) \\
&= T \sum_{z\in\mathcal{Z}} \sum_{k\in[K]} \pi_k(z) \mathbb{E}_c[\varphi(\lambda, c, k, z) \mid z]\mu(z).
\end{aligned}$$

Clearly $\pi_k(z) \in \mathcal{P}_K$, thus if we maximize over this space for the $\pi_k(z)$ instead of over $([K]^{\mathcal{Z}})^T$, it will be higher. Therefore

$$\mathrm{OPT} \leqslant T \sup_{\pi\in\mathcal{P}_K^{\mathcal{Z}}} \inf_{\lambda\in\mathbb{R}^d} \left(R^*(\lambda) + \sum_{z\in\mathcal{Z}} \sum_{k\in[K]} \pi_k(z) \mathbb{E}_c[\varphi(\lambda, c, k, z) \mid z]\mu(z) \right). \tag{3.43}$$

Notice that $\mu(z)$ plays an important role in determining the optimal policy. If we had simply run the previous algorithm for each context, this would be akin to considering that the z_t are distributed uniformly according to $\mu(z) = 1/Z$, which would have yielded a sub-optimal policy π.

Now let us take care of the lower bound. As done in eq. (3.25), we have that

$$\mathbb{E}[f_t(x_t) - p_{k_t} - \bar{R}(\gamma_t)] \geqslant \mathbb{E}[\varphi(\lambda_t, k_t, z_t, c_{tk_t}) + R^*(\lambda_t) + \langle \lambda_t, \delta_t - \gamma_t \rangle],$$

and as we did before we will use bandits algorithm on these rewards.

We would now like to bound the following adversarial contextual bandit problem regret:

$$\mathrm{Reg} = \sum_{z \in \mathcal{Z}} \max_{k \in [K]} \sum_{\substack{t=1 \\ z_t = z}}^{T} \mathbb{E}[\varphi(\lambda_t, c_t, k, z) \mid \mathcal{H}_{t-1}, z_t] - \sum_{t=1}^{T} \mathbb{E}[\varphi(\lambda_t, c_t, k_t, z_t) \mid \mathcal{H}_{t-1}, z_t].$$

The weighted estimator $\widehat{\varphi}$ is unbiased as $\varphi(\lambda_t, c_t, k_t, z_t)\mathbb{1}[k_t = k]/\pi_{tk}(z_t)$ is still unbiased. Indeed, c_t and k_t are conditionally independent given $(z_t, \mathcal{H}_{t-1})$, and $\pi_{tk}(z_t)$ is $(z_t, \mathcal{H}_{t-1})$ measurable. Thus

$$\begin{aligned}
\mathbb{E}[\frac{\varphi(\lambda_t, c_t, k_t, z_t)\mathbb{1}[k_t = k]}{\pi_{tk}(z_t)} \mid \mathcal{H}_{t-1}, z_t] &= \frac{\mathbb{E}[\varphi(\lambda_t, c_t, k_t, z_t)\mathbb{1}[k_t = k] \mid \mathcal{H}_{t-1}, z_t]}{\pi_{tk}(z_t)} \\
&= \frac{\mathbb{E}[\varphi(\lambda_t, c_t, k_t, z_t) \mid \mathcal{H}_{t-1}, z_t]\mathbb{E}[\mathbb{1}[k_t = k] \mid \mathcal{H}_{t-1}, z_t]}{\pi_{tk}(z_t)} \\
&= \mathbb{E}[\varphi(\lambda_t, c_t, k_t, z_t) \mid \mathcal{H}_{t-1}, z_t].
\end{aligned}$$

As suggested section (18.1) of [LS20a] running one anytime version of $EXP3$ (using a doubling trick for instance) on each of these public contexts $z \in \mathcal{Z}$, achieves a regret bound of order $\mathcal{O}(\sqrt{K\log(K)}\sum_{z \in \mathcal{Z}} \sqrt{\sum_{t=1}^{T} \mathbb{1}[z_t = z]})$.

Taking the expectation (for z_t) of this regret bound, and using Jensen's inequality for the square root function we have a regret term of order

$$\begin{aligned}
\mathbb{E}[\mathcal{O}(\sqrt{K\log(K)}\sum_{z \in \mathcal{Z}} \sqrt{\sum_{t \in [T]} \mathbb{1}[z_t = z]})] &\leqslant \mathcal{O}(\sqrt{K\log(K)}\sum_{z \in \mathcal{Z}} \sqrt{\mathbb{E}[\sum_{t \in [T]} \mathbb{1}[z_t = z]]}) \\
&= \mathcal{O}(\sqrt{TK\log(K)}\sum_{z \in \mathcal{Z}} \sqrt{\mu(z)}). \qquad (3.44)
\end{aligned}$$

Let $(\pi^n) \in (\mathcal{P}_K^{\mathcal{Z}})^{\mathbb{N}}$ be the sequence of contextual source mixing that maximizes the upper bound for OPT in eq. (3.43). Then because the maximum of $[K]$ points is greater than any convex combination, $\pi^n(z)$ in particular, we have

$$\begin{aligned}\sum_{z\in\mathcal{Z}}\max_{k\in[K]}\sum_{\substack{t=1\\z_t=z}}^{T}\mathbb{E}[\varphi(\lambda_t,c_t,k,z)\mid\mathcal{H}_{t-1},z] &\geqslant \sum_{z\in\mathcal{Z}}\sum_{k\in[K]}\pi_k^n(z)\sum_{\substack{t=1\\z_t=z}}^{T}\mathbb{E}[\varphi(\lambda_t,c,k,z)\mid\mathcal{H}_{t-1},z]\\ &= \sum_{k\in[K]}\sum_{t=1}^{T}\pi_k^n(z_t)\mathbb{E}[\varphi(\lambda_t,c,k,z_t)\mid\mathcal{H}_{t-1},z_t].\end{aligned}$$

In expectation this yields:

$$\begin{aligned}\mathbb{E}[\sum_{z\in\mathcal{Z}}\max_{k\in[K]}\sum_{\substack{t=1\\z_t=z}}^{T}\mathbb{E}[\varphi(\lambda_t,c_t,k,z)\mid\mathcal{H}_{t-1},z]] &\geqslant \mathbb{E}[\sum_{k\in[K]}\sum_{t=1}^{T}\pi_k^n(z_t)\mathbb{E}[\varphi(\lambda_t,c,k,z_t)\mid\mathcal{H}_{t-1},z_t]]\\ &= \sum_{k\in[K]}\sum_{t=1}^{T}\mathbb{E}[\mathbb{E}[\pi_k^n(z_t)\mathbb{E}[\varphi(\lambda_t,c,k,z_t)\mid\mathcal{H}_{t-1},z_t]\mid\mathcal{H}_{t-1}]],\end{aligned}$$

where we used the linearity of the expectation, and the tower property of the expectation for the last equality.

Because z_t and $\mathcal{H}_{t-1}$ are independent, we can first use lemma 3.11 and then the freezing lemma to express the conditional expectation given $\mathcal{H}_{t-1}$ as

$$\begin{aligned}\mathbb{E}[\mathbb{E}[\pi_k^n(z_t)\mathbb{E}[\varphi(\lambda_t,c,k,z_t)\mid\mathcal{H}_{t-1},z_t]\mid\mathcal{H}_{t-1}]] &= \mathbb{E}[\mathbb{E}[\pi_k^n(z_t)\mathbb{E}_c[\varphi(\lambda_t,c,k,z)\mid z]\mid\mathcal{H}_{t-1}]]\\ &= \mathbb{E}[\sum_{z\in\mathcal{Z}}\mu(z)\pi_k^n(z_t)\mathbb{E}_c[\varphi(\lambda_t,c,k,z)\mid z]]\end{aligned}$$

Hence,

$$\mathbb{E}[\sum_{z\in\mathcal{Z}}\max_{k\in[K]}\sum_{\substack{t=1\\z_t=z}}^{T}\mathbb{E}[\varphi(\lambda_t,c_t,k,z)\mid\mathcal{H}_{t-1},z]] \geqslant \sum_{t=1}^{T}\mathbb{E}[\sum_{k\in[K]}\sum_{z\in\mathcal{Z}}\mu(z)\pi_k^n(z)\mathbb{E}_c[\varphi(\lambda_t,c,k,z)\mid z]].$$

Grouping it together with the $R^*(\lambda_t)$ terms, and taking the $\inf$ over λ we derive that

$$\begin{aligned}&\mathbb{E}\left[\sum_{t=1}^{T}R^*(\lambda_t)+\sum_{z\in\mathcal{Z}}\max_{k\in[K]}\sum_{\substack{t=1\\z_t=z}}^{T}\mathbb{E}[\varphi(\lambda_t,c_t,k,z)\mid\mathcal{H}_{t-1},z]\right]\\ &\geqslant \sum_{t=1}^{T}\mathbb{E}\left[\inf_{\lambda\in\mathbb{R}^d}\left(R^*(\lambda)+\sum_{k\in[K]}\sum_{z\in\mathcal{Z}}\mu(z)\pi_k^n(z)\mathbb{E}_c[\varphi(\lambda,c,k,z)\mid z]\right)\right]\end{aligned}$$

$$\geqslant T \inf_{\lambda \in \mathbb{R}^d} \left(R^*(\lambda) + \sum_{k \in [K]} \sum_{z \in \mathcal{Z}} \mu(z) \pi_k^n(z) \mathbb{E}_c[\varphi(\lambda, c, k, z) \mid z] \right).$$

As $n \to \infty$ and by definition of π^n, we can conclude that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^T R^*(\lambda_t) + \sum_{z \in \mathcal{Z}} \max_{k \in [K]} \sum_{\substack{t=1 \\ z_t = z}}^T \mathbb{E}[\varphi(\lambda_t, c_t, k, z) \mid \mathcal{H}_{t-1}, z] \right] \\
&= T \sup_{\pi \in \mathcal{P}_K^{\mathcal{Z}}} \inf_{\lambda \in \mathbb{R}^d} \left(R^*(\lambda) + \sum_{z \in \mathcal{Z}} \sum_{k \in [K]} \pi_k(z) \mathbb{E}_c[\varphi(\lambda, c, k, z) \mid z] \mu(z) \right) \\
&\geqslant \mathrm{OPT}\,.
\end{aligned}$$

Finally, using eq. (3.44) and the previous inequality we obtain

$$\mathbb{E}[\sum_{t=1}^T \varphi(\lambda_t, k_t, z_t, c_t) + R^*(\lambda_t)] \geqslant \mathrm{OPT} - \mathcal{O}(\sqrt{TK \log(K)} \sum_{z \in \mathcal{Z}} \sqrt{\mu(z)}). \tag{3.45}$$

The rest of the proof follows section 3.9. $\square$

Clearly, if μ has a non-zero probability only for one z, we recover the previous regret bound without public information.

3.14.2 Infinite number of contexts — proof of proposition 3.6

Lemma. *If $\mathbb{E}[a_t(x) \mid c_{tk}, z_t]$ and $\mathbb{E}[f_t(x) \mid c_{tk}, z_t]$ are Lipschitz continuous in z with respective constants L_a and L_f for all $x \in \mathcal{X}$ with respect to $\|\cdot\|_2$, then so is φ in z with Lipschitz constant $L_a + L_f$.*

Proof. Let $\lambda \in \mathbb{R}^d$, $k \in [K]$, c_{tk} the additional information, $z_1, z_2 \in \mathcal{Z}$, and $x_1, x_2 \in \mathcal{X}$ be the maximizers that respectively yields $\varphi(\lambda, c_{tk}, k, z_1)$ and $\varphi(\lambda, c_{tk}, k, z_2)$. Because x_2 is a maximizer, we have that

$$\begin{aligned}
&\varphi(\lambda, c_{tk}, k, z_1) - \varphi(\lambda, c_{tk}, k, z_2) \\
&\leqslant \mathbb{E}[f_t(x_1) \mid c_{tk}, z_1] - \mathbb{E}[f_t(x_1) \mid c_{tk}, z_1] + \langle \lambda \mid \mathbb{E}[a_t(x_1) \mid c_{tk}, z_2] - \mathbb{E}[a_t(x_1) \mid c_{tk}, z_2] \rangle \\
&\leqslant L_f \|z_1 - z_2\|_2 + L_a \|a_t(x_1)\|_2 \|z_1 - z_2\|_2 \\
&\leqslant L_f + L_a.
\end{aligned}$$

Using the same arguments because x_1 is a maximizer, we obtain the symmetric inequality. $\square$

The main idea for the discretization (that averages the rewards over the discretized bins), is that the performance of the algorithm is close to the discretized optimal, which itself is close to the continuous optimal for smoothness reasons. However here the discretization is applied not to the original total utility $\mathcal{U}$, but instead to the bandits reward part. The number of discretized bins is then tuned depending on the lipschitz constants and the space dimension r.

The proof of the proposition proposition 3.6 then directly follows from exercise 19.5 of [LS20a] applied to the contextual bandit problem with rewards φ.

Remark. *In the case when $\mathcal{Z}$ is continuous, there are no guarantees that there exists an optimal policy π^*. Indeed, if for the product topology the space $\mathcal{P}_K^{\mathcal{Z}}$ is compact (by Tychonoff's theorem), the dual function used to derive* OPT *is not continuous. Vice versa, for a topology which makes this dual function continuous, it is unlikely for this space of functions to be compact. Hence why we need to take a maximizing sequence of π_n so that the dual function of these π_n converges to the upper bound of* OPT.

4

Multi-Armed Bandits with Guaranteed Revenue per Arm

Abstract. We consider a Multi-Armed Bandit problem with *covering* constraints, where the primary goal is to ensure that each arm receives a minimum expected reward while maximizing the total cumulative reward. In this scenario, the optimal policy then belongs to some unknown *feasible set*. Unlike much of the existing literature, we do not assume the presence of a safe policy or a feasibility margin, which hinders the exclusive use of conservative approaches. Consequently, we propose and analyze an algorithm that switches between pessimism and optimism in the face of uncertainty. We prove both precise problem-dependent and problem-independent bounds, demonstrating that our algorithm achieves the best of the two approaches – depending on the presence or absence of a feasibility margin – in terms of constraint violation guarantees. Furthermore, our results indicate that playing greedily on the constraints actually outperforms pessimism when considering *long-term* violations rather than violations on a *per-round* basis.

Contents

4.1 Introduction

4.1.1 Preliminaries

The Multi-Armed Bandit (MAB) is a classical model for sequential decision-making in the face of uncertainty [LS20b]. In the standard formulation of the problem, the objective of the learner is to maximize the total sum of rewards by efficiently balancing exploration (testing different arms to learn their rewards) and exploitation (choosing arms with the highest expected rewards based on available information). While the classic MAB framework is applicable in various domains, in many real-world problems, additional constraints and considerations come into play.

Motivation Consider, for instance, an online content recommendation system [ZB16; Zon+16], aiming to maximize user engagement and satisfaction. Its success is intricately bound to the diversity of content available on the platform. To sustain this essential diversity, platforms must ensure sufficient exposure for all content creators, guaranteeing a sufficient *revenue* for their continued activity. This revenue could be, for instance, the number of times a content was played, or revenues from ads displayed with the content, which cannot be reduced to the number of times the content is suggested by the platform. In this context, a *constrained* MAB problem (see e.g. Slivkins et al. [SSF22] and Sinha [Sin23]) emerges naturally. In this setting, the primary goal of the learner is to guarantee a fixed (known) minimum expected revenue to each arm (e.g. content recommendation) and consider the maximization of the total cumulative reward as a complementary objective.

Setting and notation We consider a K-armed bandit $\nu = (\nu_1, \dots, \nu_K) \in \mathcal{F}^K$, where $\mathcal{F}$ is a family of distributions. We denote by μ_k, the expected reward of arm k, and by $\Delta_k = \mu_1 - \mu_k$, the sub-optimality gap of arm k, assuming w.l.o.g. that arm 1 is optimal. At each time step t, the decision-maker selects an arm A_t and receives a reward $r_t \sim \nu_{A_t}$ drawn independently at random. At time t, the policy π that chooses the actions can rely on past observations $\mathcal{H}_{t-1} = (A_1, r_1, \dots, A_{t-1}, r_{t-1})$ and internal randomization. For each arm k, we denote by $N_k(t) = \sum_{s=1}^{t} \mathbb{I}(A_s = k)$, the number of times it was selected up to time t. We denote $(\cdot)_+ = \max(\cdot, 0)$, $a \wedge b = \min(a, b)$, $a \vee b = \max(a, b)$.

The goal of the decision maker is to maximize its rewards $\mathbb{E}_{\nu,\pi}\left[\sum_{t=T}^{K} r_t\right]$ under the constraint that the expected revenue of each arm scales linearly in T, namely,

$$\forall k \in [K] : \; \mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^{T} r_t \mathbb{I}(A_t = k)\right] \geqslant \lambda_k T \; , \tag{4.1}$$

where the scaling parameters $(\lambda_k)_{k\in[K]} \in (\mathbb{R}^+)^K$ are known a priori. If $\mathbb{E}_\nu[r_t | A_t = k] = \mu_k > 0$, satisfying the constraint of arm k is equivalent to ensuring that

$$\frac{1}{T}\mathbb{E}_{\nu,\pi}[N_k(T)] \geqslant p_k^\star := \frac{\lambda_k}{\mu_k}.$$

Conveniently, the previous constrained optimization problem can be reduced to the linear program

$$\min_{p \in \triangle_K} \quad \sum_{k=1}^{K} \Delta_k (p_k^\star - p_k) \text{ s.t.} \quad \forall k \in [K], p_k \geqslant p_k^\star \; ,$$

where $\triangle_K$ denotes the K-dimensional simplex. Clearly, an optimal solution consists of playing each arm with probability $p_k^\star$ (if possible) and allocating the remaining probability over the optimal arms. Unfortunately, a non-anticipative policy cannot compute $(p_k^\star)_{k\in[K]}$ or know the optimal arms because the means are unknown. Moreover, the feasibility of this linear program depends on the problem parameters, as there exists a solution if and only if $\forall k,\ \mu_k \geqslant \lambda_k$ and $\sum_{k=1}^K p_k^\star \leqslant 1$.

Definition 4.1 (Feasibility gap)**.** The feasibility gap of a problem with parameters $(\lambda, \mu) \in (\mathbb{R}^{+K})^2$ is

$$\rho_\lambda(\mu) = 1 - \sum_{k=1}^{K} \frac{\lambda_k}{\mu_k} . \tag{4.2}$$

A problem instance (λ, μ) is **feasible** if $\rho_\lambda(\mu) \geqslant 0$.

In the following, we simply denote the feasibility gap by ρ_λ when the context is clear.

Evaluation We do not assume any prior knowledge on $p_k^\star$, so relevant metrics must tolerate constraint violation to a certain level. Inspired by the literature on *safe bandits* (see next section for details), we consider two criteria to evaluate a policy on a given problem: the *excess regret* (for the regret-minimization objective), and the *constraint violation*. Denoting by $p_{k,t} = \mathbb{E}[\mathbb{I}(A_t = k)|\mathcal{H}_{t-1}]$ the *sampling probability* of arm k, we consider the following metrics.

Definition 4.2. The total per-round excess-regret and constraint violation are respectively defined by

$$\mathcal{R}_T^\pi(\nu, \lambda) = \sum_{k=1}^{K} \Delta_k \mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^{T} (p_{k,t} - p_k^\star)_+\right], \text{ and}$$
$$\mathcal{V}_T^\pi(\nu, \lambda) = \sum_{k=1}^{K} \mu_k \mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^{T} (p_k^\star - p_{k,t})_+\right] .$$

When the context is clear we omit (π, ν, λ) in the notation for simplicity. Intuitively, the per-round metrics encourage policies that smoothly converge to an optimal stationary policy. This is a desirable feature in real systems, which motivates providing policies with strong guarantees under these metrics.

4.1.2 Comparison with the Literature

Due to the numerous possible applications, the general problem of online learning with constraints covers several active research areas. In the literature, constraints typically originate from safety, fairness, or budget considerations to name but a few.

The generic problem that we consider is part of the literature on *Bandits with Linear Constraints*, notably including knapsacks (or packing) and covering constraints. Bandits with knapsacks have been extensively studied in the stochastic setting with finitely many arms [BKS18b] as well as in the contextual [AD16b] and adversarial [Imm+22] setting. Logarithmic problem-dependent bounds have also surfaced in Sankararaman and Slivkins [SS20], Li et al. [LSY21], and Kumar and Kleinberg [KK22]. Generally, positive costs are incurred and the algorithm runs until some positive threshold is violated. On the contrary, covering constraints necessitates managing negative budgets and costs. A line of works considers deterministic covering [Cla+20; Pat+21; Wan+21; Che+19], ensuring that each arm is pulled at least at a minimal known frequency, or Liu et al. [Liu+22] with deterministic linear constraints. The core setting of this paper, which is a covering problem, is more challenging because the constraints are stochastic. Some works tackle this case [AD19; SSF22], and obtain $\mathcal{O}(\sqrt{T})$ constraint violation that holds for the setting that we consider. However in [SSF22; Chz+23], this guarantee holds only with at least one strong assumption: knowledge of an initially *safe* policy[1] or a feasibility margin ($\rho_\lambda \geqslant \delta$ for some $\delta > 0$). Finally, Sinha [Sin23] studies the same revenue guarantees as described in (4.1), and propose the `BanditQ` algorithm, that implements the natural idea of sampling the arm that is the most "late" w.r.t. its revenue constraint at the current time step, up to additional mechanisms to simultaneously minimize the regret. They obtain bounds of order $\mathcal{O}(T^{3/4})$ for a long-term evaluation of constraint violation and regret, weaker than Definition 4.2.

Other areas of research are also closely related. For instance, in safe bandits [AAT19; Mor+21; Pac+21; Liu+21b; ZJ22; CGS22; HTA23] the goal is to only play actions belonging to an unknown feasibility set, with the objective of guaranteeing no violation of this constraint with high probability. The per-round evaluation metrics (def. 4.2) are inspired by some of these works. Their common approach is *pessimism-optimism* (PO): the algorithm plays the action maximizing reward (optimism) into a set included w.h.p. into the feasible set (pessimism). However, a safe-action and a known feasibility gap are again instrumental to design these algorithms. In contrast,

[1] In the current setting, this would consist in knowing for any arm k an allocation $\widetilde{p}_k \geqslant p_k^\star$ satisfying $\sum_k \widetilde{p}_k \leqslant 1$

Chen et al. [CGS23] and Agrawal and Devanur [AD19] obtained $\mathcal{O}(\sqrt{T})$ per-round safety violation without these assumptions, with a *doubly-optimistic* (DO) approach, considering instead an extended feasible set at each round. This good performance motivate the use of (DO) as a "worst-case" policy in the switching policies presented in Section 4.2. Furthermore, the algorithm `DOC` is the K-armed instance of (DO), for which we present a tighter analysis tailored for MAB.

Finally, we mention additional related fields. In online convex optimization with long term constraints [MTY09; JHA15; YNW17; Cas+22b] the learner receives full feedback of rewards and constraints. The question of unknown constraints is also considered in Chaudhary and Kalathil [CK21] and Liang et al. [LLZ23], but as done in the safe bandits literature some safe action is assumed known. This pre-existing safe policy assumption is also often made in safe Reinforcement Learning [Bur+21; Din+20; XLL20; EMP20], which additionally mainly study problem-independent guarantees. Repeated auctions with ROI constraints [CCK23; Den+23] is also similar in spirit, but again indirectly assumes some null action. Finally we mention Carlsson et al. [Car+23] that considers a setting similar to ours but with a best-policy identification objective.

In this work, we study problem dependent-bounds for stochastic bandits with unknown specific covering constraints, when no feasible actions is known beforehand and when per-round constraint violations are measured, and propose algorithms that switch between optimism and pessimism to minimize constraint violation for any possible problem.

4.1.3 Outline and Contributions

We propose several algorithms to solve the MAB problem with revenue guarantees presented in Section 4.1.1, which we frame as *Revenue-Guaranteeing Bandits* (`RGB`). `RGB` decouples the two objectives (Definition 4.2) by using a *target allocation* to satisfy the constraints, and a standard *optimistic* bandit algorithm for regret minimization. It is hence inspired by doubly-optimistic (DO) and pessimistic-optimistic (PO) approaches in the literature.

Typically, to obtain strong guarantees for the per-round constraints violations, the first natural idea is to take inspiration from (PO) methods [Pac+21; LSY21] even though there are no initial feasible actions (according to the (PO) formulation). However, for problems with a small feasibility gap, this produces poorly performing or unfeasible algorithms.

Hence we first propose and analyze DOC (Algorithm 6), based on (DO), which achieves $\mathcal{O}(\sqrt{T})$ constraint violation for all problems We refine this result with novel problem-dependent bounds (Theorem 4.4) and further prove *constant* excess regret $\mathcal{O}(\sum_k (p_k^\star \Delta_k^2)^{-1})$ if $\min_k \lambda_k > 0$.

In order to get even better problem-dependent bounds for $\mathcal{V}_T$, we introduce a new algorithm, named SPOC (Algorithm 7), built on a hybrid combination of (PO) and (DO). This policy also achieves $\mathcal{O}(\sqrt{T})$ violation for all problems, and it even gets *constant* constraints violations $\mathcal{O}(\rho_\lambda^{-1})$ (Theorem 4.5) on strictly feasible problem instances, where ρ_λ is the *feasibility gap* (Definition 4.1). SPOC thus achieves the **best of the two approaches in terms of constraint violations**. This is illustrated in Table 4.1, summarizing problem-dependent results. We only include the scaling of first-order terms and omit logarithmic factors for clarity.

We additionally prove a lower bound (Theorem 4.8) which implies that the upper bound derived for $\mathcal{V}_T$ (resp. $\mathcal{R}_T$) for DOC (resp. SPOC) cannot be improved by more than logarithmic factors.

Finally, our experiments (Section 4.4.1) suggest investigating the long-term properties of a *greedy* algorithm, named SGOC – which is thus in-between optimism and pessimism. For SGOC, we prove that the cumulative (and not the average !) *long-term* constraints violation converges to 0 (Theorem 4.9) for a phase-based version of this approach.

Tab. 4.1: Problem-dependent bounds (Section 4.3), contribution of arm k

ALG.	$\mathcal{V}_T$	$\mathcal{R}_T$
DOC	$\sqrt{p_k^\star T}$	$\frac{\mathbf{1}}{\boldsymbol{p_k^\star \Delta_k}} \wedge \frac{\mathbf{log}(\boldsymbol{T})}{\boldsymbol{\Delta_k}}$
SPOC	$\boldsymbol{\rho_\lambda^{-1}} \wedge \sqrt{\boldsymbol{p_k^\star T}}$	$\frac{1}{\mu_k}\sqrt{p_k^\star T}$
SGOC	$\sqrt{p_k^\star T}$	$\frac{1}{\mu_k}\sqrt{p_k^\star T}$

4.2 Algorithms

In this section, we detail the *revenue-guaranteeing bandit* (RGB) framework, as well as specific implementations. As detailed in Section 4.1.2, our inspiration comes from (PO) and (DO) approaches that have been proposed in the literature. In the K-armed problem considered, an RGB policy implements these principles as follows:

at each time step, a *target allocation* routine proposes a feasible target allocation $\widehat{p}_{k,t}$, where $\widehat{p}_{k,t}$ aims at estimating the allocation $p_k^\star$, while a standard *bandit algorithm* (that we call *base bandit*) chooses (independently) one arm to allocate the remaining probability $1 - \sum_{j\in[K]} \widehat{p}_{j,t}$. An arm is then chosen at random from the mixture. We detail RGB in Algorithm 5 below.

Algorithm 5 Revenue-Guaranteeing Bandits (RGB)

Input: Constraint levels $\lambda_1, \dots, \lambda_K$; algorithms `TargetAlloc` and `BaseBandit`
Init: $\mathcal{H}_0 = \{\}$ ▷ History of observations
for $t = 1, 2, \dots$ **do**
  $(\widehat{p}_{1,t}, \dots, \widehat{p}_{K,t}) = \texttt{TargetAlloc}(\mathcal{H}_{t-1})$
  $k_t = \texttt{BaseBandit}(\mathcal{H}_{t-1})$
  $p_{k,t} = \widehat{p}_{k,t} + \mathbb{1}(k_t = k)\left(1 - \sum_{j=1}^K \widehat{p}_{j,t}\right)$ for all $k \in [K]$
  Draw $A_t \sim \text{Mult}(p_{1,t}, \dots, p_{K,t})$
  Collect reward $r_t \sim \nu_{A_t}$
  Update $\mathcal{H}_t = \mathcal{H}_{t-1} \cup \{(A_t, r_t)\}$
end for

We now explore possible choices for the base bandit algorithm and for the target allocation.

Base bandit As the revenue constraints are already handled by the target allocation, the goal of the base bandit is simply to try playing rewarding arms. This can be handled by a classic bandit algorithm, which could be chosen among any standard policy, like UCB [ACF02], ETC [Per+16] TS [Tho33; AG12], KL-UCB [Cap+13] or MED [HT11; BSH23]. As detailed in the next section, choosing an optimistic algorithm is convenient for the analysis of RGB policies, so in the rest of the paper we assume that the base bandit is UCB [ACF02] or KL-UCB [Cap+13], instantiated to achieve logarithmic regret on the family of distributions $\mathcal{F}$. In the rest of the paper, we denote this algorithm by $\overline{\texttt{UCB}}$ to avoid any ambiguity with the target allocation.

We further remark that if $\max_k \lambda_k = 0$ (no revenue guarantee) then our algorithm simply follows the recommendation of $\overline{\texttt{UCB}}$, so RGB naturally interpolates standard MABs and revenue-guaranteeing bandits.

Target allocation All the target allocations considered in this paper are of the form

$$\widehat{p}^\pi_{k,t} = \frac{\lambda_k}{\widehat{\mu}^\pi_{k,t}} , \tag{4.3}$$

where for all $k \in [K]$, $\widehat{\mu}_{k,t}$ is an estimate of the mean μ_k computed with the $N_k(t-1)$ observations obtained up to time $t-1$. With a slight abuse, we say that $(\widehat{p}^\pi_{k,t})_{k\in[K]}$

is a *feasible* (resp. unfeasible) allocation if $\sum_k \widehat{p}^{\pi}_{k,t} \leqslant 1$ (resp. $\geqslant 1$). We consider $(\widehat{\mu}_{k,t})_{k\in[K],t\in[T]}$ that are either empirical means or confidence bounds. To design the latter, we use the following standard assumption.

Assumption 4.3 (Sub-Gaussian rewards). $\mathcal{F}$ is the family of 1-sub-Gaussian distributions.

Under the sub-Gaussian model, for some parameter $c > 0$ we can use the following mean estimates:

- $\widehat{\mu}^{\text{Greedy}}_{k,t} = \bar{\mu}_{k,t} := \frac{1}{N_k(t-1)} \sum_{s=1}^{t-1} r_s \mathbb{I}(A_s = k)$.
- $\widehat{\mu}^{\texttt{LCB}}_{k,t} = \texttt{LCB}_{k,t} := \bar{\mu}_{k,t} - \sqrt{\frac{6(1+c)\log(t)}{N_k(t-1)}}$.
- $\widehat{\mu}^{\texttt{UCB}}_{k,t} = \texttt{UCB}_{k,t} := \bar{\mu}_{k,t} + \sqrt{\frac{6(1+c)\log(t)}{N_k(t-1)}}$.

so that $\mathbb{P}\left(\mu_k \in [\texttt{LCB}_{k,t}, \texttt{UCB}_{k,t}]\right) \geqslant 1 - 2t^{-2(1+c)}$ (using a simple union bound on $N_k(t-1)$). As detailed in the next section, this confidence level (with $2(1+c) > 2$) is crucial for the theoretical analysis of `RGB` policies. Note however that $\overline{\texttt{UCB}}$ can use a different confidence bound than $\texttt{UCB}_{k,t}$ (e.g., with lower confidence). In the following, we use the shorthand formulation "LCB allocation" (resp. UCB or greedy) to refer to the target allocation corresponding to this estimate.

Algorithms We now detail *Doubly-Optimistic Covering* (`DOC`), *Safe Pessimistic-Optimistic Covering* (`SPOC`) and *Safe Greedy-Optimistic Covering* (`SGOC`).

It is known [AD19; CGS23] that (DO) can provide surprisingly better candidates than (PO) to satisfy the revenue guarantees when the problem has low feasibility gap. This stems from the following property:

If the problem is feasible, then the UCB allocation is feasible with high probability.

On the contrary, LCB allocations may take a long time to become feasible when ρ_λ is small (and may never be if $\rho_\lambda = 0$), which causes the failure of (PO) for low feasibility gaps. This motivates the idea of using the UCB allocation as a backup policy when the LCB one is unfeasible, which we later exploit with `SPOC`. Before that, we detail `DOC` in Algorithm 6. For completeness, `DOC` needs to be able to provide a target allocation also when the UCB one is unfeasible, even if this situation is unlikely. In the following implementation, we simply assume that a routine `UnfeasAlloc` is chosen beforehand to tackle that case. In our code we simply choose to normalize the UCB allocation if it is unfeasible, $\texttt{UnfeasAlloc}(\mathcal{H}_{t-1}) \propto \widehat{p}_{k,t}^{\texttt{UCB}}$, because dividing all revenues by the same factor seems to be a fair way to tackle unfeasibility. We discuss other choices at the end of this section.

Algorithm 6 Doubly-Optimistic Covering (DOC)

Input: $\lambda = (\lambda_1, \dots, \lambda_K)$, UnfeasAlloc
Play $\texttt{RGB}(\lambda, \texttt{UCB-Alloc}, \overline{\texttt{UCB}})$ (Alg. 5), with
UCB-Alloc:

$$\mathcal{H}_{t-1} \to \begin{cases} \left(\frac{\lambda_k}{\texttt{UCB}_{k,t}}\right)_{k\in[K]} & \text{if feasible,} \\ \texttt{UnfeasAlloc}(\mathcal{H}_{t-1}) & \text{otherwise.} \end{cases}$$

We then detail the implementation of SPOC in Algorithm 7 below. The idea is to play the LCB allocation whenever it is feasible, and to switch to the UCB allocation otherwise. We build SGOC on the same design as SPOC, replacing the LCB allocation with the greedy one. We report its implementation in Algorithm 8. Interestingly, we obtain from Equation (4.3) that for each $k \in [K], t \in [T]$, given the same observations,the sampling probabilities of the three algorithms satisfy

$$\widehat{p}_{k,t}^{\texttt{DOC}} \leqslant \widehat{p}_{k,t}^{\texttt{SGOC}} \quad \text{and} \quad \widehat{p}_{k,t}^{\texttt{DOC}} \leqslant \widehat{p}_{k,t}^{\texttt{SPOC}} \,, \tag{4.4}$$

so SGOC and SPOC are expected to serve the constraint at least as well as DOC in any situation (omitting the role of $\overline{\texttt{UCB}}$ for simplicity). This is the motivation for qualifying SPOC (resp. SGOC) as a “safe” way to implement a pessimistic (resp. greedy) approach in the RGB framework.

Algorithm 7 Safe Pessimistic-Optimistic Covering (SPOC)

Input: $\lambda = (\lambda_1, \dots, \lambda_K)$, UnfeasAlloc
Play $\texttt{RGB}(\lambda, \texttt{SPOC-Alloc}, \overline{\texttt{UCB}})$ (Alg. 5), with
SPOC-Alloc:

$$\mathcal{H}_{t-1} \to \begin{cases} \left(\frac{\lambda_k}{\texttt{LCB}_{k,t}}\right)_{k\in[K]} & \text{if feasible, else} \\ \left(\frac{\lambda_k}{\texttt{UCB}_{k,t}}\right)_{k\in[K]} & \text{if feasible,} \\ \texttt{UnfeasAlloc}(\mathcal{H}_{t-1}) & \text{otherwise.} \end{cases}$$

Remark (Individual switches). In practice, we can implement a variant of SPOC that switches as few arms as possible to the UCB allocation, in order to guarantee $\widehat{p}_{k,t} \geqslant p_k^\star$ w.h.p. for as many arms as possible. We chose the implementation of Algorithm 7 to simplify the presentation, but the theoretical guarantees derived for

SPOC in Theorem 4.5 (next section) trivially hold for any more subtle implementation of switches.

Policy for the unfeasible case In practice, the decision-maker should decide in advance what strategy to adopt if the initial problem appears to be unfeasible (which is true w.h.p. if the UCB allocation is unfeasible). For instance, for recommendation systems, there may be no way to certify in advance that some content may work "well enough" or not. However, there is no unique way to define this new goal, i.e., a new target allocation $(\widetilde{p}^\star_k)_{k\in[K]}$. This depends on the exact context of the problem, and UnfeasAlloc should be tailored to reach the chosen objective. One, that we choose in our implementations, is to avoid discriminating between arms (e.g., for fairness reasons) by defining $\widetilde{p}^\star_k \propto p^\star_k$: every arm receives the same fraction of their initial guaranteed revenue. Another is to define an implicit ranking of the arms $(i_1, \dots, i_K)$ that can be learned (by knowing a rule set depending on problem parameters) and to serve the constraints of the better-ranked arms in priority. For instance, ranking the arms by decreasing expectation minimizes the total constraint violation, while ranking the arms by increasing values of $p^\star_k$ maximizes the number of constraints that are satisfied.

Algorithm 8 Safe Greedy-Optimistic Covering (SGOC)

Input: $\lambda = (\lambda_1, \dots, \lambda_K)$, UnfeasAlloc
Play $\texttt{RGB}(\lambda, \texttt{SGOC-Alloc}, \overline{\texttt{UCB}})$ (Alg. 5), with
SGOC-Alloc:

$$\mathcal{H}_{t-1} \to \begin{cases} \left(\frac{\lambda_k}{\bar{\mu}_{k,t}}\right)_{k\in[K]} & \text{if feasible, else} \\ \left(\frac{\lambda_k}{\texttt{UCB}_{k,t}}\right)_{k\in[K]} & \text{if feasible,} \\ \texttt{UnfeasAlloc}(\mathcal{H}_{t-1}) & \text{otherwise.} \end{cases}$$

4.3 Theoretical Results

In this section, we provide upper bounds on $\mathcal{R}_T$ and $\mathcal{V}_T$ for DOC, SPOC and SGOC, and provide lower bounds that exhibit the trade-off between the two metrics. We assume that **all problems considered are feasible** (Definition 4.1), which is essential for the interpretation of the results.

4.3.1 Auxiliary Results

Before presenting the main theorems, we define the key quantities used in their statement.

Regret due to $\overline{\texttt{UCB}}$ For arms with positive revenue guarantees, it is clear that $\overline{\texttt{UCB}}$ benefits from the plays of sub-optimal arms caused by the target allocation. In Lemma 4.11 (Appendix 4.5) we prove that $\overline{\texttt{UCB}}$ only selects such arms a finite number of times. More precisely, we show that under Assumption 4.3 there exists a constant multiplicative factor α such that the number of selection of any arm $k \in [K]$ by $\overline{\texttt{UCB}}$ is upper bounded by

$$\overline{N}_k(T) := \overline{N}_k^\star \wedge \alpha \frac{\log(T)}{\Delta_k^2} \wedge T \ , \tag{4.5}$$

with a constant $\overline{N}_k^\star \in \mathbb{R} \cup \{+\infty\}$ satisfying

$$\overline{N}_k^\star = \mathcal{O}\left(\frac{\log\left(3 \vee (p_k^\star \Delta_k^2)^{-1}\right)}{p_k^\star \Delta_k^2}\right) . \tag{4.6}$$

Equation (4.5) further exhibits different regimes according to the time horizon: if $p_k^\star \leqslant (\log(T))^{-1}$ we can use the standard logarithmic bound, which is intuitive.

Sufficient sampling We now introduce a crucial result for the derivation of the problem-dependent bounds presented in the next section. It formalizes the intuition that, when playing a revenue-guaranteeing algorithm, at each step t any arm k should satisfy $N_k(t) = \Omega(p_k^\star t)$ with high probability. More specifically, we show in Lemma 4.12 (Appendix 4.5) that there exists a constant Γ_k such that if a policy π satisfies $p_{k,t}^\pi \geqslant p_{k,t}^{\texttt{DOC}}$ (which is the case for SPOC and SGOC), it holds that

$$\sum_{t=1}^{+\infty} \mathbb{P}_\pi\left(N_k(t) \leqslant \frac{p_k^\star}{8} t\right) \leqslant \Gamma_k \ , \text{ with} \tag{4.7}$$

$$\Gamma_k = \mathcal{O}\left(\frac{\log\left((p_k^\star \mu_k^3)^{-1}\right)^{\frac{3}{2}}}{p_k^\star \mu_k^3} \vee \frac{-\log(p_k^\star)}{p_k^\star} \vee \frac{1}{c}\right) . \tag{4.8}$$

The proof of this result is non-trivial and relies on a variant of Freedman's inequality (Theorem 1 from Beygelzimer et al. [Bey+11]). It is also noteworthy that the factor c^{-1} in (4.8) justifies the confidence level adopted in Algorithms 6 and 7: a smaller level may not guarantee that the arms are sufficiently sampled with high probability.

4.3.2 Upper Bounds on $\mathcal{V}_T$ and $\mathcal{R}_T$

We can now formalize our main results. When unspecified, the guarantees are problem-dependent, while problem-independent results will be explicitly stated as such. We start with DOC, which serves as a basis for the other algorithms.

Theorem 4.4 (Upper bounds for DOC)**.** *Under Assumption 4.3, the excess-regret of DOC satisfies*

$$\mathcal{R}_T^{\texttt{DOC}} \leqslant \sum_{k=1}^{K} \Big(\Delta_k(\overline{N}_k^\star + \Gamma_k)\Big) \wedge \frac{\alpha \log(T)}{\Delta_k} + \frac{K \max_k \Delta_k}{1+c} \;,$$

where α*,* $\overline{N}_k^\star$ *and* Γ_k *are respectively defined in Equations* (4.5)*,* (4.6) *and* (4.7)*. If* $\max_k \Delta_k \leqslant \Delta^+$ *for a fixed* $\Delta^+ \in \mathbb{R}$ *it furthermore holds that* $\mathcal{R}_T^{\texttt{DOC}} = \mathcal{O}(\sqrt{KT\log(T)})$ *(pb. independent bound).*

Moreover, there exists an absolute constant C_0 *such that the constraint violation DOC satisfies*

$$\mathcal{V}_T^{\texttt{DOC}} \leqslant C_0 \sum_{k=1}^{K} \sqrt{p_k^\star T \log(T)} + \sum_{k=1}^{K} \lambda_k \Gamma_k + \frac{K \max_k \mu_k}{1+c} \;,$$

and $\mathcal{V}_T^{\texttt{DOC}} = \mathcal{O}(K\sqrt{T\log(T)})$ *(pb. independent).*

The details of the proof can be found in Appendix 4.6.1. Theorem 4.4 first establishes that both $\mathcal{R}_T^{\texttt{DOC}}$ and $\mathcal{V}_T^{\texttt{DOC}}$ admit a problem-independent bound scaling in $\mathcal{O}(\sqrt{T})$, which is on par with the best results obtained in the literature for (DO) approaches (see e.g. [CGS23]). These results are refined with novel problem-dependent bounds: we obtain a constant for $\mathcal{R}_T^{\texttt{DOC}}$, and a bound for $\mathcal{V}_T^{\texttt{DOC}}$ that improves the scaling of the first-order term. For instance, if $\mu_k \gg \lambda_k$ then $\sqrt{p_k^\star T \log(T)}$ and $\lambda_k \Gamma_k = \mathcal{O}(\mu_k^{-2})$ can both be much smaller than $\sqrt{T}$. It is also noteworthy that we employed a different proof scheme to derive the two results for $\mathcal{V}_T^{\texttt{DOC}}$.

We now present upper bounds for SPOC, that we prove in Appendix 4.6.2.

Theorem 4.5 (Upper bounds for SPOC). *Under Assumption 4.3, SPOC satisfies $\mathcal{V}_T^{\texttt{SPOC}} \leqslant \mathcal{V}_T^{\texttt{DOC}}$ as well as*

$$\mathcal{V}_T^{\texttt{SPOC}} = \mathcal{O}\left(\frac{\sqrt{\log\left(\rho_\lambda^{-2} \vee e\right)}}{\rho_\lambda}\sqrt{KD_{\lambda,\mu}}\right),$$

where $D_{\lambda,\mu} = \max_{j\in[K]:\lambda_j>0} \frac{\log\left(e\vee(\lambda_j\mu_j)^{-1}\right)}{\lambda_j\mu_j}$. Moreover, there exists an absolute constant $C_1 > 0$ such that

$$\mathcal{R}_T^{\texttt{SPOC}} \leqslant \sum_{k=1}^{K} \Delta_k \left(C_1 \frac{\sqrt{p_k^\star T \log(T)}}{\mu_k} + \overline{N}_k(T) + 2\Gamma_k\right),$$

where $\overline{N}_k(T)$ and Γ_k are resp. defined in (4.5) *and* (4.7).

The main result is that $\mathcal{V}_T^{\texttt{SPOC}}$ admits a constant upper bound as soon as $\rho_\lambda > 0$, while simultaneously guaranteeing no more constraints violation than DOC for any horizon T. This justifies calling SPOC a "safe" implementation of pessimism-optimism for the revenue-guaranteeing problem. However, we note that the constant bound may be vacuous if one of the revenue parameters is very small (high constant $D_{\lambda,\mu}$). This effect may be reduced by implementing more subtle switches (see Remark 4.2), but improving $D_{\lambda,\mu}$ seems quite intricate in general.

Symmetrically to $\mathcal{V}_T^{\texttt{DOC}}$, the dominant term of the excess-regret upper bound scales as $\mathcal{O}(\sqrt{T})$. However, the factor μ_k^{-1} does not permit to obtain problem-independent results: the upper bound is vacuous for $\mu_k \leqslant T^{-1/2}$. Still, by assuming that the problem is feasible we know that $\mu_k^{-1} \leqslant \lambda_k^{-1}$, which in turn provides an upper bound on $\mathcal{R}_T^{\texttt{SPOC}}$ that is known by the decision-maker.

We recall that the assumption that the problem is feasible makes the UCB allocation feasible w.h.p., which explains why the bounds of the two theorems hold for any choice of `UnfeasAlloc`.

Finally, theoretical guarantees for SGOC can be easily derived from the bounds on $\mathcal{R}_T^{\texttt{SPOC}}$ and $\mathcal{V}_T^{\texttt{DOC}}$. We detail this in Section 4.4.2.

Remark. A minor modification of SPOC leads to problem-independent bounds on $\mathcal{R}_T^{\texttt{SPOC}}$: choose thresholds $(\tau_t)_{t\in[T]}$ such that SPOC plays the UCB allocation for arm k if $\texttt{LCB}_{k,t} \leqslant \tau_t$. With this mechanism, we can get $\mathcal{R}_T^{\texttt{SPOC}} = \mathcal{O}\left(\tau_T^{-1}\sqrt{KT} \vee \tau_T^{-3}\right)$ (up to logarithms). The drawback is that it is necessary to wait that $\tau_T \leqslant \mu_k$ to play LCB, which degrades the guarantees on $\mathcal{V}_T$ from a non-asymptotic perspective.

4.3.3 Lower Bounds

In this section we prove lower bounds that show that the problem-dependent bounds obtained in previous section for $\mathcal{V}_T^{\texttt{DOC}}$ and $\mathcal{R}_T^{\texttt{SPOC}}$ cannot be improved by more than logarithmic factors, that can depend on T but not on the problem constants. We define a revenue-guaranteeing problem by $(\lambda, \nu) \in \mathbb{R}^K \times \mathcal{F}^K$, and use the notation $\mathcal{C}$ to denote the set of feasible problems, and by $\mathcal{C}^0 \subset \mathcal{C}$ its interior (problems with positive feasibility gaps).

Definition 4.6 (Admissible policies)**.** A policy π belongs to the set of admissible policies Π if

$$\forall (\nu, \lambda) \in \mathcal{C} : \liminf_{T\to\infty} \inf_{k\in[K]} \frac{\mathbb{E}_{\nu,\pi}[N_k(T)]}{p_k^\star(\nu,\lambda)T} \geqslant 1,$$

where for $k \in [K]$, $p_k^\star(\nu, \lambda) = \frac{\lambda_k}{\mathbb{E}_{X\sim\nu_k}[X]}$.

In other words, π is admissible if all revenue guarantees are satisfied asymptotically. We now consider more precisely two subsets of admissible policies.,

Definition 4.7. A policy $\pi \in \Pi$ can be

$$\begin{aligned}
&\mathcal{R}\text{-targeting: if } \forall (\nu, \lambda) \in \mathcal{C} : \ \limsup T^{-\frac{1}{2}} \mathcal{R}_{T,\nu}^{\pi} = 0,\\
&\mathcal{V}\text{-targeting: if } \forall (\nu, \lambda) \in \mathcal{C}^0 : \ \limsup T^{-\frac{1}{2}} \mathcal{V}_{T,\nu}^{\pi} = 0 \,.
\end{aligned}$$

We denote by Π_R (resp. Π_V) the set of $\mathcal{R}$-targeting (resp. $\mathcal{V}$-targeting) policies.

It is clear from Theorem 4.4 and 4.5 that DOC is $\mathcal{R}$-targeting while SPOC is $\mathcal{V}$-targeting. We now state our main result for this part.

Theorem 4.8 (Lower bounds)**.** *Consider $\lambda \in \mathbb{R}^K$ and a bandit $\nu \in \mathcal{F}^K$ with means $(\mu_k)_{k\in[K]}$. For any policy $\pi \in \Pi_R$, it holds that*

$$\forall (\nu, \lambda) \in \mathcal{C}, \ \limsup_{T\to\infty} \frac{\mathcal{V}_T^{\pi}(\nu,\lambda)}{\sqrt{Tp_k^\star(\nu,\lambda)}} \geqslant \frac{1}{2\sqrt{e}} \,,$$

and, for any policy $\pi \in \Pi_V$ it holds that

$$\forall (\nu, \lambda) \in \mathcal{C}^0, \ \limsup_{T\to\infty} \frac{\mathcal{R}_T^{\pi}(\nu,\lambda)}{\frac{1}{\mu_k}\sqrt{Tp_k^\star(\nu,\lambda)}} \geqslant \frac{1}{2\sqrt{e}} \,.$$

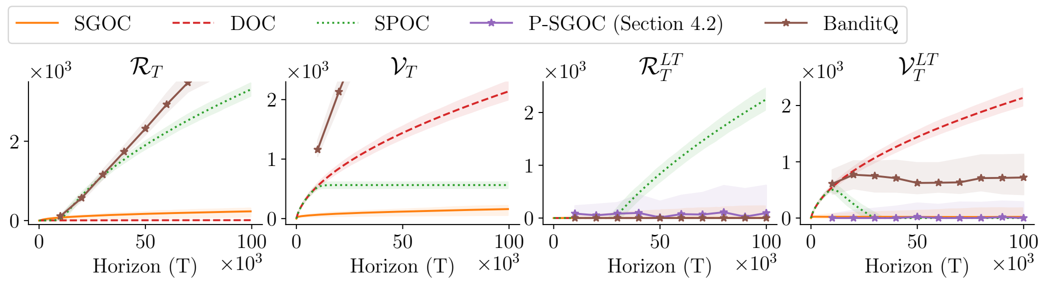


Fig. 4.1: Reproducing the simulation setup from Sinha [Sin23]

The proof details can be found in Appendix 4.6.3, and rely on standard change-of-measure arguments. Theorem 4.8 indicates that a targeting policy must pay at least $\sqrt{T}$ for the "non-targeted" objective. We furthermore observe that the upper bounds on $\mathcal{V}_T^{\texttt{DOC}}$ and $\mathcal{R}_T^{\texttt{SPOC}}$ obtained in Section 4.3.2 match the lower bounds of Theorem 4.8 up to logarithmic factors that do not depend on (λ, ν). Finally, Theorem 4.8 also confirms that the factors μ_k^{-1} in the upper bounds on $\mathcal{R}_T$ are unavoidable for $\mathcal{V}$-targeting policies.

4.4 Practical Results

4.4.1 Experiments

The experiments can be reproduced using the code, available online[2]. As highlighted in Section 4.1.2, only a small fraction of the literature is directly applicable to our setting. Thus, we benchmark `DOC`, `SPOC` and `SGOC` in terms of excess-regret and constraint violation with `BanditQ` [Sin23] and a primal-dual algorithm by Slivkins et al. [SSF22]. However, we present our results for the latter only in Appendix 4.7, since we did not obtain good performance with this algorithm. Additionally, for fair comparison with `BanditQ` we also consider the following long term metrics:

$$\mathcal{R}^{\mathrm{LT}}_{\pi,T}(\nu,\lambda) = \sum_{k=1}^{K} \Delta_k \mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^{T}(p_{k,t} - p_k^\star)\right]_+ \text{, and}$$
$$\mathcal{V}^{\mathrm{LT}}_{\pi,T}(\nu,\lambda) = \sum_{k=1}^{K} \mu_k \mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^{T}(p_k^\star - p_{k,t})\right]_+ .$$

[2] `https://github.com/DBaudry/Revenue_guaranteeing_bandits`

We replicate the experiment presented in [Sin23][3], using 200 seeds and with T varying in $[10^2, 10^5]$, reporting 10 values for non-anytime algorithms. We compute the excess regret and constraints violation as well as their long-term counterparts and display the result in Figure 4.1, averaging the values across seeds. Error bars represent the first and the last decile.

As predicted by our analysis, `DOC` has small excess regret and square root violation while `SPOC` exhibits constant violation and square root excess regret. `SGOC` exhibits $\mathcal{O}(\sqrt{T})$ excess regret and violation (see Figure 4.3 for better resolution) but still achieves lower excess regret than `SPOC` and lower constraints violation than `DOC`. In this example, we also observe the transition of `SPOC` from optimism to pessimism, making the long-term violation converge to 0. In Appendix 4.7, we further study the impact of ρ_λ on the performance of the algorithms, confirming at the same time our previous observations.

If we consider more specifically the long-term metrics, `BanditQ` seems to converge to 0 regret and to constant violation. However, simulations with different problem parameters available in Section 4.7 show that `BanditQ` exhibits positive regret on some instances, and positive violation on others, contrasting with the predictable behaviour of `SPOC` and `DOC`.

With long term metrics, `SGOC` seems to be the go-to approach, reaching both very small long-term regret and violation. It is not clear that these quantities still scales in $\sqrt{T}$. This observation motivates a closer investigation of the performance of `SGOC` w.r.t. the long term metrics. Unfortunately, providing a tight analysis for `SGOC` may be intricate, because the mean estimates $(\bar{\mu}_k(t))_{k\in[K]}$ are not independent of the trajectory. For this reason, we introduce a phase based algorithm called `P-SGOC`, in order to mimic the long-term behavior of `SGOC`. We describe and analyze this algorithm in the next section. Observe that `P-SGOC` seems to follow closely `SGOC` in the presented experiment.

4.4.2 Greedy Algorithms

We start by elaborating on the theoretical performance of `SGOC`. It is clear that, given the same sequence of observations, it holds that $p^{\texttt{DOC}}_{k,t} \leqslant p^{\texttt{SGOC}}_{k,t} \leqslant p^{\texttt{SPOC}}_{k,t}$. Hence, by following closely the proofs we can show that the upper bounds obtained for $\mathcal{V}^{\texttt{DOC}}_T$ (Theorem 4.4) hold for $\mathcal{V}^{\texttt{SGOC}}_T$, and similarly that the upper bound obtained for $\mathcal{R}^{\texttt{SPOC}}_T$ (Theorem 4.5) hold for $\mathcal{R}^{\texttt{SGOC}}_T$. This leads to the results presented in Table 4.1.

[3] $K = 5$, $\mu = (0.335, 0.203, 0.241, 0.781, 0.617)$ and $\boldsymbol{\lambda} = (0.167, 0.067, 0, 0, 0)$

Although the scaling in T and the problem parameters $(\lambda_k, \mu_k)_{k\in[K]}$ is the same, smaller multiplicative constants than C_0 and C_1 presented in the two theorems can be obtained. This is simply because the proof can involve tighter confidence intervals (since the greedy estimates do not use confidence bounds). This is on par with the simulations presented in the previous section. We now focus on the long-term metrics, providing a refined analsis for an algorithm inspired by `SGOC`.

`SGOC` with phases In this part, we assume that the horizon T is known, and that the means are bounded by 1 for simplicity[4]. `P-SGOC` proceeds in four phases. The first two phases are used to build a target allocation $\widehat{p} = (\widehat{p}_k)_{k\in[K]}$. The first phase, of length $\sum \frac{\lambda_k}{4} T$, provides a rough estimate of $(p_k^\star)$, that is then used to calibrate the length of the second estimation phase. In the third phase, `P-SGOC` serves the constraints by ensuring that $\widehat{p}_k T$ samples are collected from each arm k. Finally, $\overline{\texttt{UCB}}$ plays for the remaining time steps. We provide a detailed pseudo-code in Section 4.6.4 (Algorithm 9). In the exact implementation of `P-SGOC`, we carefully tune the length of phase 1 and 2 to ensure that the algorithm goes to phase 3 with high probability. Furthermore, using two estimation phases allows `P-SGOC` to obtain (again, w.h.p.) an uncertainty on the estimate of $p_k^\star$ on par with `SGOC` (for which the errors depends on $N_k(t) = \Omega(p_k^\star t)$ w.h.p.). For these reasons, we believe that `P-SGOC` is a good proxy for `SGOC`.

We now present the long-term guarantees of `P-SGOC`, assuming for simplicity a positive feasibility gap and only positive revenue guarantee.

Theorem 4.9 (Long-term excess-regret and constraint violation of `P-SGOC`). *Assume that* $\min_{k\in[K]} \lambda_k > 0$*, that* $\rho_\lambda > 0$*, and that* $\max_{k\in[K]} \mu_k \leqslant 1$*. If the distributions are* σ*-sub-Gaussian then* `P-SGOC` *satisfies*

$$\limsup_{T\to\infty} \mathcal{R}_T^{LT} \leqslant 24 \sum_{k=1}^{K} \frac{\sigma^2}{\mu_k^2} \Delta_k \ , \ \textit{and} \quad \limsup_{T\to\infty} \mathcal{V}_T^{LT} \leqslant 0 \ .$$

The results are stated in an asymptotic formulation to simplify their interpretation. The detailed proof, with explicit bounds, is available in Section 4.6.4.

By Theorem 4.9, `P-SGOC` asymptotically satisfies all the long-term constraints in expectation, and achieves constant excess-regret. This result is of course much stronger than what we obtained for the per-round metrics with `SGOC`, proving that a "greedy-optimistic" have merits if long-term goals are also considered.

[4] otherwise, a short preliminary phase can provide a crude upper bound on each mean.

Conclusion

In this paper we tackle a Multi-Armed Bandit problem with guaranteed per-arm revenue. Setting the *per-round* satisfaction of the revenue constraint as the main goal encourages the design of policies that switch between *pessimism* and *optimism* for the constraint estimation. This approach achieves strong theoretical guarantees, even for difficult problems with small *feasibility gap*. Numerical experiments support these findings, and further reveals the strong long-term performance of a *greedy* approach. Further theoretical results indicate that greedy outperforms pessimism under new metrics defined for *long-term* satisfaction of the constraints, by achieving constant regret and no violation.

In future works, extensions of MAB with revenue guarantees could be considered. The contextual setting [SSF22] is a natural extension but challenging as it is currently unknown whether a (DO) algorithm can get both $\mathcal{O}(\sqrt{T})$ excess regret and constraint violation. Another promising direction would be to extend the hybrid approach presented in this paper to handle more complicated constraint structures. For instance, a legal contract may specify the policies that should be targeted, whether some constraints are feasible or not, or depending on some problem parameters. We could then design algorithms with multiple switches depending on their own evaluation of their location on the decision tree.

4.5 Technical Lemmas

In this section we formalize and prove the results presented at the beginning of Section 4.3 on the regret caused by the base bandit (Lemma 4.11) and the sufficient sampling of each arm (Lemma 4.12). We also provide in Lemma 4.14 a simple result used to derive the constants presented in the two previously introduced lemmas.

In order to assess the generality of the approaches presented in this paper, we prove the aforementioned result under the following Assumption 4.10, more general than Assumption 4.3.

Assumption 4.10 (Confidence sets)**.** For any $c > 0$ and collected data $\mathcal{H}_{t-1}$, the target allocation can use a confidence interval $[\mathtt{LCB}_{k,t}, \mathtt{UCB}_{k,t}]$ satisfying

$$\mathbb{P}\left(\mu_k \in [\mathtt{LCB}_{k,t}, \mathtt{UCB}_{k,t}]\right) \geqslant 1 - \frac{1}{t^{2(1+c)}} ,$$

Furthermore, there exists a constant $C > 0$ such that

$$\texttt{UCB}_{k,t} - \texttt{LCB}_{k,t} \leqslant C\sqrt{\frac{\log(t)}{N_k(t-1)}} \,. \tag{4.9}$$

Indeed, Assumption 4.10 is not only satisfied by sub-Gaussian distributions, but by more general exponential families of distributions, and can also be applied to some families of heavy-tail distributions by building the confidence intervals with appropriate robust estimators (see e.g. [BCL13]).

Lemma 4.11 (Excess-regret caused by $\overline{\texttt{UCB}}$)**.** *We assume that the confidence bound used by $\overline{UCB}$ satisfy Assumption 4.10, and use the notation $\bar{p} = (\bar{p}_k)_{k\in[K]}$ for some arbitrary $\bar{p}_k \geqslant 0$. Then, for any time step t we denote by $\overline{UCB}_{j,t}$ the upper confidence bound used for arm j, and define $\mathcal{B}_t = \{\forall j \in [K] : \mu_j \leqslant \overline{UCB}_{j,t}\}$ (optimism), and $\mathcal{N}_{k,t} = \{N_k(t) \geqslant \bar{p}_k t\}$ (sufficient sampling). Then, the number of pulls of any sub-optimal arm k by $\overline{UCB}$ under $\mathcal{B}_t$ and $\mathcal{N}_{k,t-1}$ satisfies*

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}(A_{t+1} = k_t = k, \mathcal{B}_t, \mathcal{N}_{k,t-1})\right] \leqslant \overline{N}_k(T) := \frac{3C^2 \log\left(3 \vee \frac{C^2}{\bar{p}_k\Delta_k^2}\right)}{\bar{p}_k\Delta_k^2} \wedge C^2\frac{\log(T)}{\Delta_k^2} \wedge T \,.$$

Furthermore, for large enough time horizons this bound becomes $\overline{N}_k^\star = \frac{3C^2 \log\left(3\vee\frac{C^2}{\bar{p}_k\Delta_k^2}\right)}{\bar{p}_k\Delta_k^2}$ if $\bar{p}_k > 0$, and $C^2\frac{\log(T)}{\Delta_k^2}$ otherwise.

Proof. Let us start by fixing a sub-optimal arm k and considering $\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}(k_t = k, \mathcal{B}_t, \mathcal{D}_t, \mathcal{N}_{k,t-1})\right]$, with $\mathcal{D}_t = \{A_{t+1} = k_t\}$. First, the upper bound by T is trivial. Then, because the base bandit is based on the UCB principle we have that if $k_t = k$ then arm k has the largest upper confidence bound among all arms. In particular, if $\mathcal{B}_t$ holds then $\overline{\texttt{UCB}}_{k,t} \geqslant \overline{\texttt{UCB}}_{1,t} \geqslant \mu_1$. We thus obtain that

$$\begin{aligned}
\{k_t = k, \overline{\texttt{UCB}}_{k,t} \geqslant \max_j \overline{\texttt{UCB}}_{j,t}, \mathcal{D}_t, \mathcal{B}_t, \mathcal{N}_{k,t-1}\} &\subset \{k_t = k, \mathcal{D}_t, \overline{\texttt{UCB}}_{k,t} \geqslant \overline{\texttt{UCB}}_{1,t}, \mathcal{B}_t, \mathcal{N}_{k,t-1}\} \\
&\subset \{k_t = k, \mathcal{D}_t, \overline{\texttt{UCB}}_{k,t} \geqslant \mu_1, \mathcal{B}_t, \mathcal{N}_{k,t-1}\} \\
&\subset \left\{k_t = k, \mathcal{D}_t, \mu_k + C\sqrt{\frac{\log(t)}{N_k(t-1)}} \geqslant \mu_1, \mathcal{B}_t, \mathcal{N}_{k,t-1}\right\} \\
&= \left\{k_t = k, \mathcal{D}_t, C\sqrt{\frac{\log(t)}{N_k(t-1)}} \geqslant \Delta_k, \mathcal{B}_t, \mathcal{N}_{k,t-1}\right\} ,
\end{aligned}$$

where we used Assumption 4.10 to upper bound $\overline{\mathtt{UCB}}_{k,t} - \mu_k$. It is clear that the final event cannot happen if $N_k(t-1) \geqslant C^2 \frac{\log(t)}{\Delta_k^2}$. Hence, $C^2 \frac{\log(T)}{\Delta_k^2}$, provides the second upper bound on the number of pulls due to the base bandit, independently of the value of $\bar{p}_k$. This result is standard for the analysis of UCB algorithms. Under $\mathcal{N}_{k,t-1}$, we can further prove that the base bandit will not cause pulls arm k after a large enough time. More precisely,

$$t > t_k(\bar{p}_k) := \sup \left\{ t \in \mathbb{N} : \ \bar{p}_k t \leqslant \frac{C^2}{\Delta_k^2} \log(t) \right\} \Rightarrow \mathbb{I}(k_t = k, \mathcal{B}_t, \mathcal{D}_t, \mathcal{N}_{k,t-1}) = 0 \ .$$

Furthermore, using Lemma 4.14 we obtain that $t_k(\bar{p}_k) \leqslant \frac{3C^2 \log\left(3 \vee \frac{C^2}{\bar{p}_k \Delta_k^2}\right)}{\bar{p}_k \Delta_k^2}$, giving the first part of Lemma 4.11. This concludes the proof. $\square$

Lemma 4.12 (Sufficient sampling). *Under Assumption 4.10, for any arm $k \in [K]$ there exists a problem-dependent constant Γ_k such that if a policy π satisfies $p_{k,t}^{\pi} \geqslant p_{k,t}^{\mathtt{DOC}}$ at all time it holds that*

$$\sum_{t=1}^{+\infty} \mathbb{P}_{\pi}\left(N_k(t) \leqslant \frac{p_k^\star}{8} t\right) \leqslant \Gamma_k = \mathcal{O}\left(\frac{\log\left(\frac{1}{p_k^\star \mu_k^3}\right)^{\frac{3}{2}}}{p_k^\star \mu_k^3} \vee \frac{-\log(p_k^\star)}{p_k^\star} \vee \frac{1}{c} \right) .$$

Proof. The proof is based on the following concentration result, that proves that the sample size of each arm k is "close" to the sum of sampling probabilities with high probability.

Lemma 4.13 (Application of Freedman's inequality). *For any $\delta_t > 0$ and $\eta \in (0,1]$ it holds that*

$$N_k(t) \geqslant (1-\eta) \sum_{s=1}^{t} p_{k,s} - \frac{1}{\eta} \log(1/\delta_t) \quad \textit{with probability at least } 1 - \delta_t \ . \tag{4.10}$$

Proof. We apply Theorem 1 from [Bey+11] with the martingale difference $(X_s)_{s \leqslant t}$ defined by $\forall s \leqslant t$, $X_s = \mathbb{I}(A_s = k) - p_{k,s}$, using that $\mathbb{E}[X_s^2 | \mathcal{F}_{s-1}] = p_{k,s}(1 - p_{k,s}) \leqslant p_{k,s} \leqslant 1$. $\square$

Using Lemma 4.13 with $\delta_t = t^{-(1+p_k^\star)}$ and a parameter $\eta \in (0,1)$ (defined later) we first obtain that

$$\begin{aligned}\mathbb{P}\left(N_k(t) \leqslant \frac{p_k^\star t}{8}\right) &\leqslant \mathbb{P}\left(N_k(t) \leqslant (1-\eta)\sum_{s=1}^t p_{k,s} - \frac{1}{\eta}\log(1/\delta_t)\right) + \mathbb{P}\left((1-\eta)\sum_{s=1}^t p_{k,s} - \frac{1}{\eta}\log(1/\delta_t) \leqslant \frac{p_k^\star}{8}t\right) \\ &\leqslant \frac{1}{t^{1+p_k^\star}} + \underbrace{\mathbb{P}\left((1-\eta)\sum_{s=1}^t p_{k,s} - \frac{1+p_k^\star}{\eta}\log(t) \leqslant \frac{p_k^\star}{8}t\right)}_{P_t} .\end{aligned}$$

We then upper bound P_t by analyzing different scenarios for the sequence $(p_{k,s})_{s \leqslant t}$. We consider two events: $\mathcal{E}_t = \{\cap_{s\in[t/4,t]}\mathcal{B}_s\}$ (μ_k belongs to the confidence intervals for all rounds between $t/4$ and t), and the event $\{N_k(t/2) \geqslant n_t := \frac{C^2\log(t)}{\mu_k^2}\}$. We then obtain that

$$\begin{aligned}P_t \leqslant &\underbrace{\mathbb{P}\left((1-\eta)\sum_{s=1}^t p_{k,s} - \frac{1+p_k^\star}{\eta}\log(t) \leqslant \frac{p_k^\star}{8}t, N_k\left(\frac{t}{2}\right) \geqslant n_t, \mathcal{E}_t\right)}_{P_t'} \\ &+ \mathbb{P}\left(N_k\left(\frac{t}{2}\right) \leqslant n_t\right) + \mathbb{P}\left(\bar{\mathcal{E}}_t\right)\end{aligned}$$

We first upper bound $\sum_{t=1}^t \mathbb{P}(\bar{\mathcal{E}}_t)$ thanks to our assumptions on $\mathbb{P}(\bar{\mathcal{B}}_s)$,

$$\sum_{t=1}^T \mathbb{P}(\bar{\mathcal{E}}_t) \leqslant \sum_{t=1}^T \sum_{s=t/4}^t \mathbb{P}\left(\bar{\mathcal{B}}_s\right) \leqslant 12\sum_{t=1}^{+\infty}\frac{1}{t^{1+c}} \leqslant \frac{12}{c} , \tag{4.11}$$

We then upper bound P_t'. We first remark that

$$\mathcal{E}_t \cap \{N_k(t/2) \geqslant n_t\} \Rightarrow \forall s \in [t/2, t],\ p_{k,s} \geqslant \frac{\lambda_k}{\mu_k + C\sqrt{\frac{\log(s)}{N_k(s)}}} \geqslant \frac{p_k^\star}{2} . \tag{4.12}$$

Using this result, we obtain the following deterministic upper bound

$$\begin{aligned}\sum_{t=1}^T P_t' &\leqslant \sum_{t=1}^T \mathbb{I}\left(\frac{p_k^\star t}{4}(1-\eta) - \frac{1+p_k^\star}{\eta}\log(t) \leqslant \frac{p_k^\star t}{8}\right) \\ &\leqslant \sum_{t=1}^T \mathbb{I}\left(\frac{p_k^\star t}{8} - \left(\frac{p_k^\star t}{4}\eta + \frac{1+p_k^\star}{\eta}\log(t)\right) \leqslant 0\right)\end{aligned}$$

We now optimize η to make this quantity as small as possible, obtaining $\eta = 2\sqrt{\frac{(1+p_k^\star)\log(t)}{p_k^\star t}}$. We can use this value only if $\eta < 1$, but remark that this check is redundant with the indicator being 0. Indeed, we obtain that

$$\begin{aligned}\sum_{t=1}^{T} P'_t &\leqslant \sum_{t=1}^{T} \mathbb{I}\left(\frac{p_k^\star t}{8} \leqslant \sqrt{p_k^\star(1+p_k^\star)t\log(t)} \cup 2\sqrt{\frac{(1+p_k^\star)\log(t)}{p_k^\star t}} > 1\right) \\ &\leqslant \sum_{t=1}^{T} \mathbb{I}\left(\frac{t}{\log(t)} \leqslant 64\left(1+\frac{1}{p_k^\star}\right)\right) := t_k^0 \ .\end{aligned}$$

Using Lemma 4.14 we further obtain that $t_k^0 = \mathcal{O}\left(\frac{-\log(p_k^\star)}{p_k^\star}\right)$.

It remains to upper bound $\sum_{t=1}^{T} \mathbb{P}\left(N_k(t/2) \leqslant n_t\right)$. We can use the exact same scheme as before, with the significant advantage that n_t scales in $\log(t)$ instead of t, so we can use a cruder lower bound on $\sum_{s=1}^{t/2} p_{k,s}$ to conclude the proof. We now define $\varepsilon'_t = \{\cap_{s\in[t/4,t/2]}\mathcal{B}_s\}$, which provides

$$\varepsilon'_t \Rightarrow \sum_{s=1}^{t/2} p_{k,s} \geqslant \frac{t}{4} \times \frac{\lambda_k}{\mu_k + C\sqrt{\log(t)}} \ .$$

Using this result along with Lemma 4.13 for $\delta'_t = \frac{1}{t^{1+\lambda_k}}$ we now obtain that for any $\eta \in (0,1)$,

$$\mathbb{P}\left(N_k\left(\frac{t}{2}\right) \leqslant n_t\right) \leqslant \delta'_t + \mathbb{P}\left((1-\eta)\sum_{s=1}^{t} p_{k,s} - \frac{1+\lambda_k}{\eta}\log(t) \leqslant n_t, \mathcal{E}'_t\right) + \mathbb{P}(\bar{\mathcal{E}}'_t) \ .$$

Similarly as for $\sum \mathbb{P}(\bar{\mathcal{E}}_t)$ we obtain that $\sum_{t=1}^{T} \mathbb{P}(\bar{\mathcal{E}}'_t) \leqslant \frac{8}{c}$, and we also obtain by choosing $\eta = \sqrt{8\frac{(1+\lambda_k)\log(t)}{\lambda_k t}(\mu_k + C\sqrt{\log(t)})}$ in Lemma 4.13 that

$$\begin{aligned}&\sum_{t=1}^{T} \mathbb{P}\left((1-\eta)\sum_{s=1}^{t} p_{k,s} - \frac{1+\lambda_k}{\eta}\log(t) \leqslant n_t, \mathcal{E}'_t\right) \\ &\leqslant \sum_{t=1}^{T} \mathbb{I}\left((1-\eta)\frac{\lambda_k}{\mu_k + C\sqrt{\log(t)}}\frac{t}{8} - \frac{1+\lambda_k}{\eta}\log(t) \leqslant n_t\right) \\ &= \sum_{t=1}^{T} \mathbb{I}\left(\frac{\lambda_k}{\mu_k + C\sqrt{\log(t)}}\frac{t}{8} - \sqrt{\frac{\lambda_k(1+\lambda_k)t\log(t)}{2(\mu_k + C\sqrt{\log(t)})}} \leqslant n_t\right) \\ &\leqslant \sum_{t=1}^{T} \mathbb{I}\left(\frac{\lambda_k}{\mu_k + C\sqrt{\log(t)}}\frac{t}{8} \leqslant 2n_t,\right) \vee \sum_{t=1}^{T} \mathbb{I}\left(\frac{\lambda_k}{\mu_k + C\sqrt{\log(t)}}\frac{t}{8} \leqslant 2\sqrt{\frac{\lambda_k(1+\lambda_k)t\log(t)}{2(\mu_k + C\sqrt{\log(t)})}}\right) \\ &:= t_k^1 \vee t_k^2 \ ,\end{aligned}$$

that are both problem-dependent constants. Similarly to t_k^0, an upper bound on t_k^1 and t_k^2 can be derived explicitly thanks to Lemma 4.14, and again we used that the value of η that we choose is valid for $t \geqslant t_k^1 \vee t_k^2$. We thus easily obtain that $t_k^1 = \mathcal{O}\left(\frac{1}{\lambda_k \mu_k^2} \log\left(\frac{1}{\lambda_k \mu_k^2}\right)^{3/2}\right)$ and $t_k^2 = \mathcal{O}\left(\frac{1}{\lambda_k} \log\left(\frac{1}{\lambda_k}\right)^{3/2}\right)$. A summary of all the results obtained so far finally leads to

$$\sum_{t=1}^{+\infty} \mathbb{P}\left(N_k(t) \leqslant \frac{p_k^* t}{8}\right) \leqslant \Gamma_k := t_k^0 \vee t_k^1 \vee t_k^2 + \underbrace{\frac{1}{p_k^\star} + \frac{1}{\lambda_k}}_{\sum_{t=1}^{T}(\delta_t + \delta_t')} + \frac{20}{c} \,. \tag{4.13}$$

If c is not unreasonably small, the "characteristic times" $(t_k^i)_{i \in [3]}$ dominate this bound $p_k^\star$ and/or μ_k are small. Thanks to Lemma 4.14, we obtain the scaling

$$\Gamma_k = \mathcal{O}\left(\frac{1}{p_k^\star \mu_k^3} \log\left(\frac{1}{p_k^\star \mu_k^3}\right)^{3/2} \vee \frac{1}{p_k^\star \mu_k} \log\left(\frac{1}{p_k^\star \mu_k}\right)^{3/2} \vee \frac{-\log(p_k^\star)}{p_k^\star} \vee \frac{1}{c}\right) ,$$

$\square$

and remark that the second term can be removed without changing the result.

Lemma 4.14. *For any $\alpha \geqslant 1$, the mapping $f_\alpha : x \in [(\alpha+2)^\alpha \vee 3, \infty) \mapsto \sup\left\{t \in \mathbb{N} : \frac{t}{\log(t)^\alpha} \leqslant x\right\}$ satisfies*

$$f_\alpha(x) \leqslant (\alpha+2)^\alpha \times \log(x)^\alpha x.$$

Proof. We start by remarking that the function $g(x) = \frac{x}{\log(x)^\alpha}$ is strictly increasing for all $x \geq e^\alpha$. Now, consider a value $s = Ax\log(x)^\alpha$ for some $A > 0$, such that $s \geqslant 3 \vee e^\alpha$. By the monotonicity of $\frac{t}{(\log t)^\alpha}$, we have that

$$t > s \Rightarrow \frac{t}{(\log(t)^\alpha} > \frac{s}{(\log(s)^\alpha} = x \times \frac{A\log(x)^\alpha}{(\log(A) + \log(x) + \alpha\log(\log(x)))^\alpha} \,.$$

Then, for $x \geqslant A \geqslant 3$, it holds that $\log(A) + \log(x) + \alpha\log(\log(x)) \leqslant (\alpha+2)\log(x)$, so we can simply choose $A = (\alpha+2)^\alpha$ to obtain the result.

All that is left is to verify that for this choice, $s = (\alpha+2)^\alpha \times \log(x)^\alpha x \geqslant 3 \vee e^\alpha$, but this clearly holds for all $x \geq 3$ and $\alpha > 0$. $\square$

4.6 Detailed Proofs

In this section we prove all the main results presented in this paper.

4.6.1 Proof of Theorem 4.4

We recall the theorem before detailing the proof.

Theorem 4.15 (Upper bounds for DOC)**.** *Under Assumption 4.3, the excess-regret of DOC satisfies*

$$\mathcal{R}_T^{\texttt{DOC}} \leqslant \sum_{k=1}^{K} \left(\Delta_k(\overline{N}_k^\star + \Gamma_k)\right) \wedge \frac{\alpha \log(T)}{\Delta_k} + \frac{K \max_k \Delta_k}{1+c} \,,$$

where α*,* $\overline{N}_k^\star$ *and* Γ_k *are respectively defined in Equations* (4.5)*,* (4.6) *and* (4.7)*. If* $\max_k \Delta_k \leqslant \Delta^+$ *for a fixed* $\Delta^+ \in \mathbb{R}$ *it furthermore holds that* $\mathcal{R}_T^{\texttt{DOC}} = \mathcal{O}(\sqrt{KT\log(T)})$ *(pb. independent bound).*

Moreover, there exists an absolute constant C_0 *such that the constraint violation DOC satisfies*

$$\mathcal{V}_T^{\texttt{DOC}} \leqslant C_0 \sum_{k=1}^{K} \sqrt{p_k^\star T \log(T)} + \sum_{k=1}^{K} \lambda_k \Gamma_k + \frac{K \max_k \mu_k}{1+c} \,,$$

and $\mathcal{V}_T^{\texttt{DOC}} = \mathcal{O}(K\sqrt{T\log(T)})$ *(pb. independent).*

Proof. We divide the proof between the upper bounds on the excess-regret and constraint violation, starting with the excess-regret. As for the technical results of Section 4.5, we write the proof under the more general Assumption 4.10, and instantiate the constants for Assumption 4.3 only when the final results are derived.

Upper bound on the excess-regret We consider the events $\mathcal{N}_{k,t} = \{N_k(t) \geqslant \bar{p}_k t\}$ for some $(\bar{p}_k)_{k\in[K]}$, $\mathcal{B}_t = \{\forall j \in [K] : \ \mu_j \in [\texttt{LCB}_{j,t}, \texttt{UCB}_{j,t}]\}$, and $\mathcal{D}_t = \{A_{t+1} = k_t\}$.

Let us fix a sub-optimal arm k. For any time step t, we use that $p^{\texttt{DOC}}_{k,t} \leqslant \frac{\lambda_k}{\texttt{UCB}_{k,t}}$ to obtain that

$$
\begin{aligned}
&\mathbb{E}[(p^{\texttt{DOC}}_{k,t} - p^\star_k)_+] \\
&\leqslant \mathbb{E}[(p^{\texttt{DOC}}_{k,t} - p^\star_k)_+ \mathbb{I}(\mathcal{N}_{k,t-1}, \mathcal{B}_t)] + \mathbb{P}(\bar{\mathcal{N}}_{k,t-1}) + \mathbb{E}\left[p^{\texttt{DOC}}_{k,t}\mathbb{I}(\bar{\mathcal{B}}_t)\right] \\
&\leqslant \mathbb{E}\left[\left(\frac{\lambda_k}{\texttt{UCB}_{k,t}} - p^\star_k\right)_+ \mathbb{I}(\mathcal{N}_{k,t-1}, \mathcal{B}_t)\right] + \mathbb{E}\left[\mathbb{I}(k_t = k, \mathcal{N}_{k,t-1}, \mathcal{B}_t, \mathcal{D}_t)\right] + \mathbb{P}(\bar{\mathcal{N}}_{k,t-1}) + \mathbb{E}\left[p^{\texttt{DOC}}_{k,t}\mathbb{I}(\bar{\mathcal{B}}_t)\right] \\
&\leqslant 0 + \mathbb{E}\left[\mathbb{I}(k_t = k, \mathcal{N}_{k,t-1}, \mathcal{B}_t, \mathcal{D}_t)\right] + \mathbb{P}(\bar{\mathcal{N}}_{k,t-1}) + \mathbb{E}\left[p^{\texttt{DOC}}_{k,t}\mathbb{I}(\bar{\mathcal{B}}_t)\right] \ .
\end{aligned}
$$

We emphasize that due to the optimism under $\mathcal{B}_t$, this equality holds for any choice of $(\bar{p}_k)_{k\in[K]}$.

Now, we first obtain the term $\frac{K \max_k \Delta_k}{1+c}$ by upper bounding the following term,

$$
\sum_{t=1}^T \sum_{k=1}^K \Delta_k \mathbb{E}\left[p_{k,t}\mathbb{I}(\bar{\mathcal{B}}_t)\right] \leqslant \sum_{t=1}^T \max_k \Delta_k \mathbb{P}(\bar{\mathcal{B}}_t) \tag{4.14}
$$

$$
\leqslant \left(\sum_{t=1}^T \delta_t\right) K \max_k \Delta_k = \frac{K \max_k \Delta_k}{1+c} \ . \tag{4.15}
$$

We then obtain the first-order term of the result by using Lemma 4.11 and Lemma 4.12. We now consider two different choices for $(\bar{p}_k)_{k\in[K]}$ to bound the two remaining terms.

<u>Case I:</u> $\bar{p}_k = \frac{p^\star_k}{8}$. We can use both lemmas and obtain that $\forall k \in [K]$,

$$
\sum_{t=1}^T \left(\mathbb{E}\left[\mathbb{I}(k_t = k, \mathcal{N}_{k,t-1}, \mathcal{B}_t, \mathcal{D}_t)\right] + \mathbb{P}(\bar{\mathcal{N}}_{k,t-1})\right) \leqslant \overline{N}^\star_k + \Gamma_k \ ,
$$

where $\overline{N}^\star_k$ and Γ_k are respectively defined in the statement of Lemma 4.11 and of Lemma 4.12.

<u>Case II:</u> $\bar{p}_k = 0$. For this choice, $(\mathcal{N}_{k,t})$ always holds ($\mathbb{P}(\bar{\mathcal{N}}_{k,t}) = 0$), and we complete the result by using Lemma 4.11 to obtain that, at the same time, this quantity is also bounded by $\alpha\frac{\log(T)}{\Delta_k^2}$, for some $\alpha = C^2$ (with C defined in Assumption 4.10).

Moreover, this second upper bound by $\alpha\frac{\log(T)}{\Delta_k}$ also guarantees the standard $\mathcal{O}\left(\sqrt{KT\log(T)}\right)$ problem-independent bound, which is directly obtained by taking the maximum between the logarithmic bound and the trivial bound by T. We remark that the upper bound Δ^+ on the gap is necessary to upper bound the term $\frac{K \max_k \Delta_k}{1+c}$.

Constraint violation To upper bound $\mathcal{V}_T^{\texttt{DOC}}$, we consider any arm $k \in [K]$ for which $\lambda_k > 0$. We again use the events $\mathcal{N}_{k,t}$ and $\mathcal{B}_t$ that we used to upper bound $\mathcal{R}_T^{\texttt{DOC}}$, so that under $\mathcal{B}_t$ the UCB allocation is feasible. We first write that

$$\sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t}^{\texttt{DOC}})_+\right] \leqslant \sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t}^{\texttt{DOC}})_+ \mathbb{I}(\mathcal{N}_{k,t-1}, \mathcal{B}_t)\right] + \lambda_k \sum_{t=1}^{T} \mathbb{P}(\bar{\mathcal{N}}_{k,t-1}) + \mu_k \sum_{t=1}^{T} \mathbb{E}\left[p_k^* \mathbb{I}(\bar{\mathcal{B}}_t)\right] .$$

We upper bound the second order terms using Assumption 4.10 and Lemma 4.12, obtaining (similar to the proof for the excess regret) $\sum_{k=1}^{K} \lambda_k \Gamma_k + \frac{K \max_k \mu_k}{1+c}$ when summing over the k arms. For the remaining term, we use Assumption 4.10 to write that

$$\begin{aligned}
\leqslant \sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t}^{\texttt{DOC}})_+ \mathbb{I}(\mathcal{N}_{k,t-1}, \mathcal{B}_t)\right] &\leqslant \sum_{t=1}^{T} \mu_k \lambda_k \mathbb{E}\left[\left(\frac{\texttt{UCB}_{k,t} - \mu_k}{\texttt{UCB}_{k,t} \times \mu_k}\right)_+ \mathbb{I}(\mathcal{N}_{k,t-1}, \mathcal{B}_t)\right] \\
&\leqslant \sum_{t=1}^{T} p_k^\star \times C \sqrt{8 \frac{\log(t)}{p_k^\star t}} ,
\end{aligned}$$

which concludes the proof for the problem-dependent bound, with $C_0 = 2\sqrt{8}C$. Under Assumption 4.10 we can define $C = 2\sqrt{6(1+c)}$, which further provides $C_0 = 16\sqrt{3(1+c)}$.

For the problem-independent bound, we take another path that do not use the events $\mathcal{N}_{k,t}$, since the scaling of Γ_k provided by Lemma 4.12 does not allow us to recover the desired bound. We use that the value of $\texttt{UCB}_{k,t}$ is determined by $(\widehat{\mu}_{k,n})_{n \in \mathbb{N}}$ (empirical average with sample size $N_k(t-1) = n$), t and the confidence level δ_t. Since the confidence level is increasing with t and δ_t, we can claim that there exists an absolute constant D such that

$$\forall t \in T \text{ , } \texttt{UCB}_{k,t} \leqslant \widetilde{\texttt{UCB}}(N_k(t-1), T) := \mu_k + D\sqrt{\frac{\log(T)}{N_k(t-1)}} \text{ with probability at least } 1 - \frac{1}{T} .$$

We denote by $\mathcal{B}$ the corresponding good event and $\widetilde{p}_{k,t} = \frac{\lambda_k}{\widetilde{\texttt{UCB}}(N_k(t-1),T)}$, and write that

$$\begin{aligned}
\sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t}^{\texttt{DOC}})_+\right] &\leqslant \sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - \widetilde{p}_{k,t}^{\texttt{DOC}})_+\right] \\
&\sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - \widetilde{p}_{k,t}^{\texttt{DOC}})_+ \mathbb{I}(\mathcal{B})\right] + T \mu_k \mathbb{P}(\bar{\mathcal{B}}) \\
&\leqslant \sum_{t=1}^{T} \mu_k \lambda_k \mathbb{E}\left[\left(\frac{\widetilde{\texttt{UCB}}_k(N_k(t-1), T) - \mu_k}{\widetilde{\texttt{UCB}}(N_k(t-1), T) \times \mu_k}\right)_+ \mathbb{I}(\mathcal{B})\right] + \mu_k
\end{aligned}$$

$$\leqslant Dp_k^\star \times \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} \frac{\mu_k}{\widetilde{\texttt{UCB}}(N_k(t-1),T)} \sqrt{\frac{\log(T)}{N_k(t-1)}} \mathbb{I}(\mathcal{B})\right]}_{Z_T} + \mu_k \,.$$

We further upper bound Z_T using a union bound on $N_k(t-1)$,

$$Z_T \leqslant \mathbb{E}\left[\sum_{t=1}^{T}\sum_{n=1}^{T} \frac{\mu_k}{\widetilde{\texttt{UCB}}(n,T)} \sqrt{\frac{\log(T)}{n}} \mathbb{I}(N_k(t-1)=n,\mathcal{B})\right]$$
$$\sum_{n=1}^{T} \frac{\mu_k}{\widetilde{\texttt{UCB}}(n,T)} \sqrt{\frac{\log(T)}{n}} \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}(N_k(t-1)=n,\mathcal{B})\right] \,.$$

Without loss of generality, we assume that $n \geq 1$; the case of $n = 0$ is taken care of by sampling each arm once at the beginning, which would not change any of the results.

Notice that the sum in the expectation is the number of rounds since we get to $N_k(t-1) = n$ until we play arm k another time and move to $N_k(t-1) = n+1$, under $\mathcal{B}$. We use again that under $\mathcal{B}$, for all time steps t the sampling probability of k is larger than $\widetilde{p}_k(t) = \frac{\lambda_k}{\widetilde{\texttt{UCB}}_k(n,T)}$ when $N_k(t-1) = n$. Thus, $\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{I}(N_k(t-1)=n,\mathcal{B})\right]$ is smaller than the expectation of a geometric random variable with probability $\frac{\lambda_k}{\widetilde{\texttt{UCB}}_k(n,T)}$. We hence obtain that

$$Z_T \leqslant \sum_{n=1}^{T} \frac{\mu_k}{\widetilde{\texttt{UCB}}(n,T)} \sqrt{\frac{\log(T)}{n}} \times \frac{\widetilde{\texttt{UCB}}(n,T)}{\lambda_k} = \sum_{n=1}^{T} \frac{1}{p_k^\star} \sqrt{\frac{\log(T)}{n}}$$

Multiplying Z_T by $Dp_k^\star$, we then conclude that

$$\sum_{t=1}^{T} \mu_k \mathbb{E}\left[p_k^\star - p_{k,t}^{\texttt{DOC}}\right] \leqslant 2D\sqrt{\log(T)T} \,,$$

which gives the problem-independent bound of $\mathcal{O}(K\sqrt{T\log(T)}$ when summing over the K constraints. □

4.6.2 Proof of Theorem 4.5

We recall the theorem below.

Theorem 4.16 (Upper bounds for SPOC)**.** *Under Assumption 4.3, SPOC satisfies* $\mathcal{V}_T^{\textit{SPOC}} \leqslant \mathcal{V}_T^{\textit{DOC}}$ *as well as*

$$\mathcal{V}_T^{\textit{SPOC}} = \mathcal{O}\left(\frac{\sqrt{\log\left(\rho_\lambda^{-2} \vee e\right)}}{\rho_\lambda} \sqrt{K D_{\lambda,\mu}} \right) ,$$

where $D_{\lambda,\mu} = \max_{j\in[K]:\lambda_j>0} \frac{\log\left(e \vee (\lambda_j \mu_j)^{-1}\right)}{\lambda_j \mu_j}$*. Moreover, there exists an absolute constant* $C_1 > 0$ *such that*

$$\mathcal{R}_T^{\textit{SPOC}} \leqslant \sum_{k=1}^K \Delta_k \left(C_1 \frac{\sqrt{p_k^\star T \log(T)}}{\mu_k} + \overline{N}_k(T) + 2\Gamma_k \right) ,$$

where $\overline{N}_k(T)$ *and* Γ_k *are resp. defined in* (4.5) *and* (4.7).

Proof. We decompose the proof in two parts, starting with the per-round constraint violation. We again write the proof under the more general Assumption 4.10, and instantiate the constants for Assumption 4.3 only when the final results are derived.

Constraint violation: By design, SPOC uses the UCB target allocation if the LCB one is unfeasible. Hence, it directly holds that

$$\forall k \in [K]: \ p_{k,t}^{\texttt{SPOC}} \geqslant p_{k,t}^{\texttt{DOC}} \Rightarrow \mathcal{V}_T^{\texttt{SPOC}} \leqslant \mathcal{V}_T^{\texttt{DOC}} .$$

This gives a first part of the result. We now consider the refined upper bound for problems with positive feasibility gap. Assume that $\rho_\lambda > 0$, and denote by $\mathcal{G}_t = \left\{ \sum_{j=1}^K \frac{\lambda_j}{\texttt{LCB}_{j,t}} \leqslant 1 \right\}$ the event that the LCB allocation is feasible, and by $\mathcal{B}_t = \{\forall j \in [K]: \ \mu_j \in [\texttt{LCB}_{j,t}, \texttt{UCB}_{j,t}]\}$ the event that all means are well concentrated. We further define $\mathcal{N}_t = \cap_j \mathcal{N}_{j,t} \coloneqq \cap_j \{N_j(t) \geqslant p_j^\star \frac{t}{8}\}$, and first write that for any arm k with $\lambda_k > 0$ it holds that

$$\sum_{t=1}^T \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t})_+\right] \leqslant \sum_{t=1}^T \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t})_+ \mathbb{I}(\mathcal{N}_{t-1}, \mathcal{B}_t)\right] + \lambda_k \sum_{t=1}^T \mathbb{P}(\bar{\mathcal{N}}_t) + \mu_k \sum_{t=1}^T \mathbb{E}[p_k^\star \mathbb{I}(\bar{\mathcal{B}}_t)]$$

We can upper bound the last two terms similarly as in the proof of Theorem 4.4. Using Lemma 4.12 and Assumption 4.10 we obtain that

$$\sum_{k=1}^K \left(\lambda_k \sum_{t=1}^T \mathbb{P}(\bar{\mathcal{N}}_t) + \mu_k \sum_{t=1}^T \mathbb{E}[p_k^\star \mathbb{I}(\bar{\mathcal{B}}_t)] \right) \leqslant \left(\sum_{j=1}^K \lambda_j \right) \sum_{k=1}^K \Gamma_k + \frac{K \max_k \mu_k}{1+c} , \quad (4.16)$$

where we used a union bound to derive the first term. We now consider the first-order term of the result. Again, we use that by design $p_{k,t}^{\texttt{SPOC}} \geqslant p_{k,t}^{\texttt{DOC}}$ for any k, t. The result comes from using the same proof as for DOC up to a characteristic time t_ρ depending on ρ_λ and other problem parameters. We start by writing that

$$
\begin{aligned}
V_T^k := \sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t})_+ \mathbb{I}(\mathcal{N}_{t-1}, \mathcal{B}_t)\right] &\leqslant \sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - p_k^\star \mathbb{I}(\mathcal{G}_t) + p_{k,t}^{\texttt{DOC}} \mathbb{I}(\bar{\mathcal{G}}_t))_+ \mathbb{I}(\mathcal{N}_{t-1}, \mathcal{B}_t)\right] \\
&= \sum_{t=1}^{T} \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t}^{\texttt{DOC}})_+ \mathbb{I}(\mathcal{N}_{t-1}, \mathcal{B}_t, \bar{\mathcal{G}}_t)\right] .
\end{aligned}
$$

Then, we use a simple property to introduce ρ_λ in the analysis: if $\forall j \in [K]$: , $\bar{\mu}_j^{\texttt{LCB}}(t) \geqslant (1 - \rho_\lambda)\mu_j$, then the LCB allocation is feasible and $\mathcal{G}_t$ holds. Combining this property with $\mathcal{B}_t$ and $\mathcal{N}_{t-1}$, we obtain that the events considered cannot hold simultaneously when t is larger than a problem-dependent constant, formally

$$
t > t_\rho := \sup \left\{ t \in \mathbb{N} : \ \exists j \in [K] : \ C\sqrt{8\frac{\log(t)}{p_j^\star t}} \geqslant \rho_\lambda \mu_j \right\} . \tag{4.17}
$$

By Lemma 4.14, we furthermore obtain that $t_\rho \leqslant \max_{j \in [K]} \frac{24C^2}{p_j^\star (\rho_\lambda \mu_j)^2}$. We can thus write that

$$
\begin{aligned}
V_T^k &\leqslant \sum_{t=1}^{t_\rho} \mu_k \mathbb{E}\left[(p_k^\star - p_{k,t}^{\texttt{DOC}})_+ \mathbb{I}(\mathcal{N}_{t-1}, \mathcal{B}_t)\right] \\
&\leqslant C \sum_{t=1}^{t_\rho} \sqrt{\frac{8p_k^\star \log(t)}{t}} && \text{(as in Theorem 4.4)} \\
&\leqslant 2C\sqrt{8p_k^\star t_\rho \log(t_\rho)} \\
&\leqslant \frac{56C^2}{\rho_\lambda} \max_{j \in [K]} \sqrt{\frac{p_k^\star}{\lambda_j \mu_j} \log\left(\frac{112C^2}{\lambda_j \mu_j \rho_\lambda^2}\right)} .
\end{aligned}
$$

Remarking that for $(x, y) > 0$ it holds that $x + y \leqslant ((y \vee 1)(1 + x))$, we can further write that

$$
V_T^k \leqslant 56C^2 \times \frac{1 \vee \sqrt{\log\left(\frac{1}{\rho_\lambda^2}\right)}}{\rho_\lambda} \max_{j \in [K]} \sqrt{\frac{p_k^\star}{\lambda_j \mu_j}\left(1 + \log\left(\frac{112C^2}{\lambda_j \mu_j}\right)\right)}, \tag{4.18}
$$

and finally using that $\sum_k \sqrt{p_k^\star} \leqslant \sqrt{K}$ we get

$$\sum_{k=1}^{K} V_T^k \leqslant \frac{\sqrt{\log\left(\frac{1}{\rho_\lambda^2} \vee e\right)}}{\rho_\lambda} \sqrt{K D_{\lambda,\mu}}\ ,$$

where $D_{\lambda,\mu}$ is explicitly defined in Equation (4.18). This completes the proof regarding the constraint violation.

Excess regret: Let us consider again the events $\mathcal{B}_t = \{\forall j \in [K] \mu_j \in [\mathtt{LCB}_{j,t}, \mathtt{UCB}_{j,t}]\}$ and $\mathcal{N}_{k,t} = \{N_k(t) \geqslant \frac{p_k^\star}{8} t\}$, $\mathcal{D}_t = \{A_{t+1} = k_t\}$, and $\mathcal{G}_t = \left\{\sum_{k=1} p_{k,t}^{\mathtt{LCB}} \leqslant 1\right\}$. We re-use our results for the LCB allocation, remarking that

$$\begin{aligned}\mathcal{R}_T^{\mathtt{SPOC}} \leqslant & \underbrace{\sum_{k=1}^{K} \Delta_k \mathbb{E}\left[(p_{k,t}^{\mathtt{SPOC}} - p_k^\star)_+ \mathbb{I}(\mathcal{B}_t, \mathcal{N}_{k,t-1}, \mathcal{G}_t)\right]}_{\widetilde{\mathcal{R}}_T^{\mathtt{LCB}}} \\ & + \underbrace{\sum_{k=1}^{K} \Delta_k \mathbb{E}\left[(p_{k,t}^{\mathtt{DOC}} - p_k^\star)_+ \left(\mathbb{I}(\mathcal{B}_t, \mathcal{N}_{k,t-1}, \mathcal{D}_t) + \mathbb{I}(\overline{\mathcal{B}_t, \mathcal{N}_{k,t-1}})\right)\right]}_{\widetilde{\mathcal{R}}_T^{\mathtt{DOC}}}\ .\end{aligned}$$

With the same arguments as for the proof of Theorem 4.4, we first obtain that

$$\widetilde{\mathcal{R}}_T^{\mathtt{DOC}} \leqslant \sum_{k=1}^{K} \Delta_k \left(\overline{N}_k(T) + \Gamma_k\right)\ .$$

It remains to upper bound the term $\widetilde{\mathcal{R}}_T^{\mathtt{LCB}}$. We first remark that if $\rho_\lambda = 0$ then the LCB allocation is unfeasible under $\mathcal{B}_t$ so $\mathcal{R}_T^{\mathtt{LCB}} = 0$. Let us now consider the case $\rho_\lambda > 0$. Since lower bounding the number of rounds for which the LCB allocation is feasible is intricate, we simply drop the event $\mathcal{G}_t$ in the rest of the proof after using that

$$(p_{k,t}^{\mathtt{SPOC}} - p_k^\star)_+ \mathbb{I}(\mathcal{G}_t) \leqslant p_k^\star \left(\frac{\mu_k - \mathtt{LCB}_{k,t}}{\mathtt{LCB}_{k,t}}\right)_+ \vee 1\ .$$

so that we can now work with the confidence intervals. Under $\mathcal{N}_{k,t-1}$ and $\mathcal{B}_t$, we are sure that $\mathtt{LCB}_{k,t} \geqslant \frac{\mu_k}{2}$ if

$$t > t_k(1/2) := \sup\left\{s \in \mathbb{N} :\ C\sqrt{\frac{8\log(t)}{p_k^\star t}} \geqslant \frac{\mu_k}{2}\right\}\ .$$

We remark from the proof of Lemma 4.12 that this term is one of the component of Γ_k, so for simplicity we use $t_k(1/2) \leqslant \Gamma_k$ in the statement of Theorem 4.5. We hence obtain that

$$\widetilde{\mathcal{R}}_T^{\texttt{LCB}} \leqslant \sum_{k=1}^{K} \Delta_k \Gamma_k + \underbrace{\sum_{k=1}^{K} \Delta_k \sum_{t=t_k(1/2)}^{T} 2\mathbb{E}\left[p_k^\star \left(\frac{\mu_k - \texttt{LCB}_{k,t}}{\mu_k}\right) \mathbb{I}(\mathcal{B}_t, \mathcal{N}_{k,t-1})\right]}_{R_T} .$$

Finally, using Assumption 4.10 we obtain that

$$\begin{aligned} R_T &\leqslant \sum_{k=1}^{K} \Delta_k \sum_{t=t_k(1/2)}^{T} 2C \frac{p_k^\star}{\mu_k} \sqrt{\frac{8\log(t)}{p_k^\star t}} \\ &\leqslant 4C \sum_{k=1}^{K} \frac{\Delta_k}{\mu_k} \sqrt{8 p_k^\star T \log(T)} , \end{aligned}$$

which concludes the proof, obtaining that $C_1 = 4C\sqrt{8}$. Under Assumption 4.3, since $C = 2\sqrt{6(1+c)}$ we obtain that $C_1 = 32\sqrt{3(1+c)}$. $\square$

4.6.3 Proof of Theorem 4.8

We recall from Definition 4.7 that Π_R denotes the set of $\mathcal{R}$-targeting policies ($\mathcal{R}_T = o(\sqrt{T})$) and Π_V denotes the set of $\mathcal{V}$-targeting policies ($\mathcal{V}_T = o(\sqrt{T})$). We prove the following result.

Theorem 4.17 (Lower bounds). *Consider $\lambda \in \mathbb{R}^K$ and a bandit $\nu \in \mathcal{F}^K$ with means $(\mu_k)_{k\in[K]}$. For any policy $\pi \in \Pi_R$, it holds that*

$$\forall (\nu, \lambda) \in \mathcal{C},\ \limsup_{T\to\infty} \frac{\mathcal{V}_T^\pi(\nu,\lambda)}{\sqrt{T p_k^\star(\nu,\lambda)}} \geqslant \frac{1}{2\sqrt{e}} ,$$

and, for any policy $\pi \in \Pi_V$ it holds that

$$\forall (\nu, \lambda) \in \mathcal{C}^0,\ \limsup_{T\to\infty} \frac{\mathcal{R}_T^\pi(\nu,\lambda)}{\frac{1}{\mu_k}\sqrt{T p_k^\star(\nu,\lambda)}} \geqslant \frac{1}{2\sqrt{e}} .$$

Proof. Let us fix a set of parameters $(\lambda_1, \dots, \lambda_k)_{\in(\mathbb{R}^+)^K}$. We start by proving the first statement, assuming that the bandit regret under the policy $\pi \in \Pi$ is dominated by $\sqrt{T}$ asymptotically for all problems. We fix a bandit instance $\nu \in \mathcal{F}$ and, for simplicity, let us consider an arbitrary arm with constraint parameter $\lambda > 0$. In the following, we assume w.l.o.g. that the selected arm is arm 1 (up to re-indexing the arms and constraints).

Then, consider another bandit instance $\nu' \in \mathcal{F}^K$, where the distributions of the arms in (ν, ν') are the same except for arm 1. To keep simple notation, let us now denote by ν and ν' the distribution of arm 1 (only) under the two models.

Let $\mu = \mathbb{E}_{X\sim\nu}[X]$ be the expectation of this arm under ν. Then, choose $\nu' \in \mathcal{F}$ to be absolutely continuous w.r.t. ν and such that $\mathbb{E}_{X\sim\nu'}[X] = \mu + \varepsilon$ for some $\varepsilon > 0$. Assume that $(\lambda_1, \ldots, \lambda_K)$ are such that the problem is feasible under ν (by extension, it is feasible under ν'). We further denote by p_t the sampling probability of the selected arm at time step t for a trajectory under π, and use the shorthand notation $p^\star_\nu = \frac{\lambda}{\mu}$ and $p^\star_{\nu'} = \frac{\lambda}{\mu+\varepsilon}$. We consider the event

$$\mathcal{E} = \left\{ \sum_{t=1}^T p_t \leqslant \frac{\lambda}{\mu + \frac{\varepsilon}{2}} T \right\} .$$

The result follows from a standard change of measure argument. We first remark that

$$\mathcal{E} \text{ holds under } \nu \Rightarrow \sum_{t=1}^T (p^\star_\nu - p_t)_+ \geqslant \sum_{t=1}^T (p^\star_\nu - p_t) \geqslant T\lambda \frac{\varepsilon}{2\mu\left(\mu + \frac{\varepsilon}{2}\right)} ,$$

and similarly that

$$\bar{\mathcal{E}} \text{ holds under } \nu' \Rightarrow \sum_{t=1}^T (p_t - p^\star_{\nu'})_+ \geqslant \sum_{t=1}^T (p_t - p^\star_{\nu'}) \geqslant T\lambda \frac{\varepsilon}{2(\mu+\varepsilon)\left(\mu + \frac{\varepsilon}{2}\right)} .$$

We use the Bretagnolle-Huber inequality (see e.g. Theorem 14.2 of [LS20b]),

$$\begin{aligned}
&\mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^T (p^\star_\nu - p_t)_+\right] + \mathbb{E}_{\nu',\pi}\left[\sum_{t=1}^T (p_t - p^\star_{\nu'})_+\right] \\
&\geqslant \mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^T (p^\star_\nu - p_t)_+ \mathbb{I}(\mathcal{E})\right] + \mathbb{E}_{\nu',\pi}\left[\sum_{t=1}^T (p_t - p^\star_{\nu'})_+ \mathbb{I}(\bar{\mathcal{E}})\right] \\
&\geqslant T\lambda \frac{\varepsilon}{2(\mu+\varepsilon)\left(\mu + \frac{\varepsilon}{2}\right)} \left(\mathbb{P}_{\nu,\pi}(\mathcal{E}) + \mathbb{P}_{\nu',\pi}(\bar{\mathcal{E}})\right) \\
&\geqslant T\lambda \frac{\varepsilon}{2(\mu+\varepsilon)\left(\mu + \frac{\varepsilon}{2}\right)} \exp\left(-\mathbb{E}_{\nu,\pi}[N_1(T)] D(\nu, \nu')\right) .
\end{aligned}$$

By Assumption 4.3, as ν and ν' are sub-Gaussian it holds that $D(\nu, \nu') \geqslant \frac{\varepsilon^2}{2}$

$$\mathbb{E}_{\nu,\pi}\left[\sum_{t=1}^{T}(p_\nu^\star - p_t)_+\right] + \mathbb{E}_{\nu',\pi}\left[\sum_{t=1}^{T}(p_t - p_{\nu'}^\star)_+\right] \geqslant \frac{\lambda}{2\mu^2}\varepsilon T \times \underbrace{\frac{\exp\left(-\mathbb{E}_{\nu,\pi}[N_1(T)]\frac{\varepsilon^2}{2}\right)}{\left(1+\frac{\varepsilon}{\mu}\right)^2}}_{B(T,\varepsilon)}$$

We now consider the properties of $B(T,\varepsilon)$ for small values of ε and large values of T. Since π is an admissible policy (Definition 4.6), it must hold that $\liminf \mathbb{E}_{\nu,\pi}[N_1(T)] \geqslant p_\nu^\star T$. Hence, for $T \to \infty$ and $\varepsilon \to 0$ we have $\varepsilon B(T,\varepsilon) \sim \varepsilon e^{-p_\nu^\star T \frac{\varepsilon^2}{2}}$, which is maximized by choosing $\varepsilon = (p_\nu^\star T)^{-\frac{1}{2}}$. This choice provides $B(T,\varepsilon) \sim e^{-1/2}$. To complete the proof of the first lower bound, we use that π is $\mathcal{R}$-targeting so $\frac{\mathbb{E}_{\nu',\pi}\left[\sum_{t=1}^{T}(p_t - p_\nu^\star)_+\right]}{\sqrt{T}} \to 0$. Hence, the scaling in $\sqrt{T}$ must come from the contribution of arm 1 to $\mathcal{V}_T^\pi(\nu,\lambda)$. Furthermore, we obtain an asymptotic rate of $\frac{1}{2\sqrt{e}}\sqrt{Tp_\nu^\star}$.

We omit the proof of the second statement of the theorem, since it consists in the exact same steps. The only subtlety is that we now need to assume that ν satisfies $\rho_\lambda(\nu) > 0$ (as indicated in the statement), so that for $\varepsilon > 0$ small enough the relevant alternative problem ν' (such that $\mathbb{E}_{X\sim\nu_1'} = \mu_1 - \varepsilon$) can be feasible for ε small enough. Furthermore, the event of interest for this part of the proof becomes $\mathcal{E} = \left\{\sum_{t=1}^{T} p_t \geqslant \frac{\lambda}{\mu - \frac{\varepsilon}{2}} T\right\}$, so that we can lower bound the excess regret suffered by arm 1 under the assumption that π is $\mathcal{V}$-targeting. □

4.6.4 Proof of Theorem 4.9

In order to make the presentation of the proof of Theorem 4.9 clearer, we detail in Algorithm 9 the implementation of `P-SGOC`.

We then recall the theorem, before proving it.

Theorem 4.18 (Long-term excess-regret and constraint violation of `P-SGOC`). *Assume that $\min_{k\in[K]} \lambda_k > 0$, that $\rho_\lambda > 0$, and that $\max_{k\in[K]} \mu_k \leqslant 1$. If the distributions are σ-sub-Gaussian then `P-SGOC` satisfies*

$$\limsup_{T\to\infty} \mathcal{R}_T^{LT} \leqslant 24 \sum_{k=1}^{K} \frac{\sigma^2}{\mu_k^2}\Delta_k \ , \ \textit{and} \quad \limsup_{T\to\infty} \mathcal{V}_T^{LT} \leqslant 0 \ .$$

Proof. We start by defining the favorable high-probability events that guarantee the performance of the algorithm.

Algorithm 9 Phased Safe Greedy-Optimistic Covering (`P-SGOC`)

Input: $\lambda = (\lambda_1, \ldots, \lambda_K)$, time horizon T
Set $\mathcal{S}_K = \{k \in [K] : \ \lambda_k > 0\}$

Phase 1: (Initial estimation)
for $k \in \mathcal{S}_K$ **do**
  Collect $N_k^1 = \frac{\lambda_k}{4} T$ samples $(r_{k,1}, \ldots, r_{k,N_k^1})$ ▷ Total phase duration $\leqslant T/4$ if feasible
  Compute $\widehat{\mu_k^1} = \frac{1}{N_k^1} \sum_{i=1}^{N_k^1} r_{k,i}$ ▷ Crude estimate
  Store data in $\mathcal{H}_1$
end for

Phase 2: (Refined estimation)
for $k \in \mathcal{S}_K$ **do**
  if $\sum_{j=1}^{K} \left(\frac{\lambda_j}{6\widehat{\mu_j^1}} \vee \frac{\lambda_j}{2} \right) T \leqslant \frac{T}{2}$ **then**
    $N_k^2 = \left(\frac{\lambda_j}{6\widehat{\mu_j^1}} \vee \frac{\lambda_j}{2} \right) T$ ▷ Typically $N_k^2 = \frac{\lambda_k}{6\widehat{\mu_k^1}} T$
  else
    $N_k^2 = \frac{\lambda_k}{2}$ ▷ Always $N_k^2 \geqslant \frac{\lambda_k}{2}$
  end if
  Collect N_k^2 samples $(r_{k,1,2}, \ldots, r_{k,N_k^2,2})$
  Compute $\widehat{\mu_k^2} = \frac{1}{N_k^2} \sum_{i=1}^{N_k^2} r_{k,i,2}$ ▷ Refined estimate
  Store data in $\mathcal{H}_2$
end for

Phase 3: (Target allocation)
for $k \in \mathcal{S}_K$ **do**
  if $\sum_{j=1}^{K} \frac{\lambda_j}{\widehat{\mu_j^2}} \leqslant 1$ **then**
    $N_k^3 = \left(\frac{\lambda_k}{\widehat{\mu_k^2}} - \frac{N_k^1 + N_k^2}{T} \right)_+ T$ ▷ Target total $\frac{\lambda_k}{\widehat{\mu_k^2}} T$
  else
    Compute $\widehat{\mathtt{UCB}_j^2}$ such that $\mathbb{P}(\widehat{\mathtt{UCB}_j^2} \leqslant \mu_j) \leqslant \frac{1}{T}$ for all $j \in [K]$
    $N_k^3 = \left(\frac{\lambda_k}{\widehat{\mathtt{UCB}_k^2}} - \frac{N_k^1 + N_k^2}{T} \right)_+ T$ ▷ Optimistic fallback
  end if
  Collect N_k^3 samples from each arm $k \in \mathcal{S}_K$ via round-robin (stop if T is reached)
  Store data in $\mathcal{H}_3$
end for

Phase 4: (Regret minimization)
Play `UCB` using $\mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3$ until horizon T is reached ▷ Final phase with base bandit

Success events We consider events that ensure that the algorithm goes to phase 3 with a good estimate $\widehat{\mu_k^2}$. First, we use that if for an arm j the estimate $\widehat{\mu_j^1}$ is well concentrated it is possible to collect $N_j^2 = \Omega(p_j^\star T)$ samples from this arm. More precisely, we have that

$$\mathcal{G}_1 \coloneqq \left\{ \forall j \in [K] :\ \widehat{\mu_j^1} \in \left[\frac{2\mu_j}{3}, 2\mu_j\right] \right\} \Rightarrow \forall j \in [K] :\ N_j^2 \in \left[\frac{p_j^\star}{12}T, \frac{p_j^\star}{2}T\right] .$$

In particular, under $\mathcal{G}_1$ the algorithm goes to phase 3 with a number of samples for each arm that is a fraction of their optimal allocation. Furthermore, the number of pulls of arm k in the third phase is positive if $\widehat{\mu_k^2} \leqslant 2\mu_k$. Considering that we also would not want the sampling probability in the third phase to be too high, we naturally consider the second type of good events,

$$\forall k : \mathcal{G}_k^2 \coloneqq \left\{ \widehat{\mu_k^2} \in \left[\frac{1}{2}\mu_k, 2\mu_k\right] \right\} ,$$

so that under $\mathcal{G} = \mathcal{G}^1 \cap \mathcal{G}_k^2$ the algorithm collects $\frac{\lambda_k}{\widehat{\mu_k^2}}T$ samples for arm k, and $N_k^2 \geqslant \frac{p_k^\star}{12}T$.

Before upper bounding the two metrics, we first provide an auxiliary results that will be used to prove both statements. We upper bound the probability of the bad event, as follows

$$\begin{aligned}
\mathbb{P}(\bar{\mathcal{G}}) &\leqslant \sum_{j=1}^{K} \left\{ \mathbb{P}\left(\widehat{\mu_j^1} \notin \left[\frac{2\mu_j}{3}, 2\mu_j\right]\right) + \mathbb{P}\left(\widehat{\mu_k^2} \in \left[\frac{1}{2}\mu_k, 2\mu_k\right], \widehat{\mu_j^1} \in \left[\frac{2\mu_j}{3}, 2\mu_j\right]\right) \right\} \\
&\leqslant \sum_{j=1}^{K} \left\{ \mathbb{P}\left(\widehat{\mu_j^1} \notin \left[\frac{2\mu_j}{3}, 2\mu_j\right]\right) + \mathbb{P}\left(\widehat{\mu_k^2} \in \left[\frac{1}{2}\mu_k, 2\mu_k\right], N_j^2(T) \geqslant \frac{p_j^\star}{12}T\right) \right\} .
\end{aligned}$$

All the terms can then be upper bounded using Hoeffding's inequality, so we obtain that

$$\mathbb{P}(\bar{\mathcal{G}}) \leqslant \sum_j \left\{ e^{-\frac{\lambda_j T}{2\sigma_j^2}\left(\frac{2\mu_j}{3}\right)^2} + e^{-\frac{\lambda_j T}{2\sigma_j^2}(2\mu_j)^2} + e^{-\frac{p_j^\star}{24\sigma_j^2}T\left(\frac{1}{2}\mu_j\right)^2} + e^{-\frac{p_j^\star}{24\sigma_j^2}T(2\mu_j)^2} \right\} . \tag{4.19}$$

Then, before proving the statements we can also remark that (1) for the long-term metrics the quantity of interest for the analysis is $\mathbb{E}\left[\sum_{t=1}^T p_{k,t}^{\texttt{P-SGOC}}\right]$, and (2) that for deterministic algorithm we can consider the expected number of pulls of each arm instead of the sampling probabilities. Indeed, we can write a phase decomposition as follows

$$\mathbb{E}\left[\sum_{t=1}^{T} p_{k,t}^{\texttt{P-SGOC}}\right] = \mathbb{E}[N_k(T)] = \sum_{s=1}^{4} \mathbb{E}[N_k^i]$$

where N_k^i denotes the number of pulls of arm k in phase i, for $i \in [4]$ (set to 0 if P-SGOC does not reach this phase).

Upper bounding $\mathcal{R}_T^{\mathsf{LT}}$ Under $\mathcal{G}$, the algorithm goes to phase 3 and plays the greedy allocation obtained at the end of phase 2. Further using that $\mathcal{G} \subset \mathcal{G}' := \{N_k^2 \geqslant \frac{p_k^\star}{12}T, \widehat{\mu_k^2} \geqslant \mu_k/2\}$ (as explained above), we can then upper bound $\sum_{s=1}^{4} \mathbb{E}[N_k^i]$ as follows,

$$\begin{aligned}
\sum_{s=1}^{4} \mathbb{E}[N_k^i(T)] &\leqslant T\mathbb{E}\left[\frac{\lambda_k}{\widehat{\mu_k^2}} \times \mathbb{I}\left(\mathcal{G}'\right)\right] + T\mathbb{P}(\bar{\mathcal{G}}') + \mathbb{E}[N_k^4(T)\mathbb{I}(\mathcal{G}')] \\
&= \underbrace{Tp_k^\star\, \mathbb{E}\left[\frac{\mu_k}{\widehat{\mu_k^2}} \times \mathbb{I}\left(\mathcal{G}'\right)\right]}_{A_T} + T\mathbb{P}(\bar{\mathcal{G}}) + \mathbb{E}[N_k^4(T)\mathbb{I}(\mathcal{G}')] \,.
\end{aligned}$$

The upper bound of Equation (4.19) indicates that $\lim\limits_{T\to\infty} T\mathbb{P}(\bar{\mathcal{G}}) = 0$. Then, by a direct adaptation of the proof of Lemma 4.11 we can also obtain that, since $\lambda_k > 0$, $\lim\limits_{T\to\infty} \mathbb{E}[N_k^4(T)\mathbb{I}(\mathcal{G})] = 0$. Hence, for a large enough horizon T excess-regret can only be caused by the term A_T.

We then rewrite A_T in the form of a bias and variance term as follows,

$$\begin{aligned}
A_T &= \mathbb{E}\left[\frac{1}{1 + \frac{\widehat{\mu_k^2} - \mu_k}{\mu_k}} \times \mathbb{I}\left(\mathcal{G}'\right)\right] \\
&= \mathbb{E}\left[\left(1 - \frac{\widehat{\mu_k^2} - \mu_k}{\mu_k} + \frac{\left(\frac{\widehat{\mu_k^2} - \mu_k}{\mu_k}\right)^2}{1 + \frac{\widehat{\mu_k^2} - \mu_k}{\mu_k}}\right)\mathbb{I}\left(\mathcal{G}'\right)\right] \quad , \quad \text{since } \frac{1}{1+x} = 1 - x + \frac{x^2}{1+x} \text{ for } x > -1.
\end{aligned}$$

Then, we use that for any sample size n it holds that $\mathbb{E}\left[\widehat{\mu_k^2}\mathbb{I}(\widehat{\mu_k^2} \geqslant \mu_k/2, N_k^2 = n)\right] \geqslant \mu_k$ (the estimate is positively biased by $\mathcal{G}'$), so the bias term is negative. We can thus further upper bound A_T by

$$
\begin{aligned}
A_T &\leqslant 1 + \mathbb{E}\left[\frac{\left(\frac{\widehat{\mu_k^2} - \mu_k}{\mu_k}\right)^2}{1 + \frac{\widehat{\mu_k^2} - \mu_k}{\mu_k}} \mathbb{I}\left(\mathcal{G}'\right)\right] \\
&\leqslant 1 + 2\mathbb{E}\left[\left(\frac{\widehat{\mu_k^2} - \mu_k}{\mu_k}\right)^2 \mathbb{I}\left(\mathcal{G}'\right)\right], \text{ since } \widehat{\mu_k^2} \geqslant \frac{\mu_k}{2} \\
&\leqslant 1 + \frac{2}{\mu_k^2} \times \mathbb{E}\left[\left(\widehat{\mu_k^2} - \mu_k\right)^2 \mathbb{I}\left(N_k^2 \geqslant \frac{p_k^\star}{12} T\right)\right].
\end{aligned}
$$

Then, for the simplicity of notation we denote by $\bar{\mu}_{k,n}$ the mean estimate corresponding to $N_k^2 = n$. Using that the sample size and that the variance of the mean estimate are independent, we obtain that

$$
\begin{aligned}
\mathbb{E}\left[\left(\widehat{\mu_k^2} - \mu_k\right)^2 \mathbb{I}\left(N_k^2 \geqslant \frac{p_k^\star}{12} T\right)\right] &= \sum_{n=\frac{p_k^\star}{12}T}^{T} \mathbb{P}(N_k^2 = n) \mathbb{E}\left[(\bar{\mu}_{k,n} - \mu_k)^2\right] \\
&\leqslant \max_{n \geqslant \frac{p_k^\star}{12}T} \mathbb{E}\left[(\bar{\mu}_{k,n} - \mu_k)^2\right] \\
&\leqslant 12 \frac{\sigma^2}{p_k^\star T}.
\end{aligned}
$$

We then obtain the desired constant by multiplying this upper bound by $2\frac{p_k^\star T}{\mu_k^2}$, which proves the upper bound provided on $\mathcal{R}_T^{\mathrm{LT}}$: P-SGOC suffers constant regret when $\min_k \lambda_k > 0$.

Lower Bounding $\mathcal{V}_T^{\mathrm{LT}}$ We now lower bound $\mathbb{E}[N_k(T)]$, and consider the successful event $\mathcal{G}_1$. Under this event, we are sure that the duration of phase 2 is no more than $T/2$ rounds, that the third phase occurs and that the mean estimate of each arm j used in phase 3 is computed with at least $\frac{p_k^\star}{12} T$ samples. We furthermore define $\mathcal{J} = \left\{ \sum_{j=1}^{K} \frac{\lambda_j}{\mu_j^2} \leqslant 1 \right\}$ the event that the greedy allocation proposed at the end of phase 2 is feasible. In this part of the analysis, we omit phase 4 for simplicity, as well as the case when $\bar{\mathcal{J}}$ holds. Indeed, since we assume that $\rho_\lambda > 0$ and we consider the asymptotic problem-dependent bound, we can avoid considering the cases where the UCB allocation is played. Hence, we simply consider the following lower bound,

$$
\mathbb{E}[N_k(T)] \geqslant \mathbb{E}\left[\left(N_k^1 + N_k^2 + N_k^3\right) \mathbb{I}(\mathcal{G}, \mathcal{J})\right]
$$

$$\geqslant T\mathbb{E}\left[\frac{\lambda_k}{\widehat{\mu_k^2}}\mathbb{I}\left(\mathcal{J},\mathcal{G}_1\right)\right] .$$

In general, there might be a problem of definition for $\widehat{\mu_k^2}^{-1}$ (that can technically be negative or infinite, even though μ_k is assumed to be positive). However under $\mathcal{J}$ the greedy allocation is feasible, which is possible only if $\widehat{\mu_k^2} \geqslant \lambda_k \geqslant \frac{\lambda_k}{2}$. We can thus use that

$$\frac{\lambda_k}{\widehat{\mu_k^2}}\mathbb{I}\left(\mathcal{J},\mathcal{G}_1\right) = \frac{\lambda_k}{\widehat{\mu_k^2}+\left(\frac{\lambda_k}{2}-\widehat{\mu_k^2}\right)_+}\mathbb{I}\left(\mathcal{J},\mathcal{G}_1\right) ,$$

which is a convenient re-writing because the right-hand term is now well-defined even when ignoring the events $\mathcal{J}$ and $\mathcal{G}_1$. Using also that under $\mathcal{G}_1$ it holds that $N_k^2 \geqslant \frac{p_k^\star}{12}T$, we obtain that

$$\begin{aligned}
\mathbb{E}[N_k(T)] &\geqslant T\mathbb{E}\left[\frac{\lambda_k}{\widehat{\mu_k^2}+\left(\frac{\lambda_k}{2}-\widehat{\mu_k^2}\right)_+}\mathbb{I}\left(\mathcal{J},\mathcal{G}_1\right)\right] \\
&= T\mathbb{E}\left[\frac{\lambda_k}{\widehat{\mu_k^2}+\left(\frac{\lambda_k}{2}-\widehat{\mu_k^2}\right)_+}\mathbb{I}\left(\mathcal{J},\mathcal{G}_1, N_k^2 \geqslant \frac{p_k^\star}{12}T\right)\right] \\
&= T\mathbb{E}\left[\frac{\lambda_k}{\widehat{\mu_k^2}+\left(\frac{\lambda_k}{2}-\widehat{\mu_k^2}\right)_+}\mathbb{I}\left(N_k^2 \geqslant \frac{p_k^\star}{12}T\right)\left(1-\mathbb{I}\left(\overline{\mathcal{J},\mathcal{G}_1}\right)\right)\right] \\
&\geqslant \underbrace{T\mathbb{E}\left[\frac{\lambda_k}{\widehat{\mu_k^2}+\left(\frac{\lambda_k}{2}-\widehat{\mu_k^2}\right)_+}\mathbb{I}\left(N_k^2 \geqslant \frac{p_k^\star}{12}T\right)\right]}_{B_T} - 2T\left(\mathbb{P}(\bar{\mathcal{J}},\mathcal{G}_1)+\mathbb{P}(\bar{\mathcal{G}}_1)\right) .
\end{aligned}$$

We then prove that the probabilities corresponding to $\bar{\mathcal{J}}$ and $\bar{\mathcal{G}}_1$ are negligible asymptotically. First, we can upper bound $T\mathbb{P}\left(\bar{\mathcal{G}}_1\right)$ by (4.19) (only the first two terms of the r.h.s. are necessary). Then, we similarly upper bound

$$T\mathbb{P}(\bar{\mathcal{J}},\mathcal{G}_1) \leqslant T\mathbb{P}\left(\exists j:\ \lambda_j > 0 \text{ and } \widehat{\mu_j^2} \leqslant (1-\rho_\lambda)\mu_j, \mathcal{G}_1\right) \leqslant T\sum_{j=1}^{K} e^{-\frac{p_j^\star}{24\sigma^2}T(\rho_\lambda\mu_j)^2} \to 0 ,$$

which is again negligible asymptotically, because $\rho_\lambda > 0$. We can thus focus on lower bounding B_T, and use the independence of N_k^2 and $\widehat{\mu}_k^2$ to write (using the same notation $\bar{\mu}_{k,n}$ that we used when upper bounding $\mathcal{R}_T^{\mathrm{LT}}$) that

$$\begin{aligned}
B_T &= p_k^\star T \sum_{n=\frac{p_k^\star}{12}T} \mathbb{P}(N_k^2 = n)\mathbb{E}\left[\frac{\mu_k}{\bar{\mu}_{k,n} + \left(\frac{\lambda_k}{2} - \bar{\mu}_{k,n}\right)_+}\right] \\
&\geqslant p_k^\star T \sum_{n=\frac{p_k^\star}{12}T} \mathbb{P}(N_k^2 = n)\left(1 - \frac{1}{\mu_k}\mathbb{E}\left[\bar{\mu}_{k,n} - \mu_k\right] - \frac{1}{\mu_k}\mathbb{E}\left[\left(\frac{\lambda_k}{2} - \bar{\mu}_{k,n}\right)_+\right]\right) \\
&= p_k^\star T \sum_{n=\frac{p_k^\star}{12}T} \mathbb{P}(N_k^2 = n)\left(1 - \frac{1}{\mu_k}\mathbb{E}\left[\left(\frac{\lambda_k}{2} - \bar{\mu}_{k,n}\right)_+\right]\right) \\
&\geqslant p_k^\star T\left(1 - \mathbb{P}\left(N_k^2 \leqslant \frac{p_k^\star}{12}T\right)\right)\left(1 - \frac{p_k^\star}{2}e^{-\frac{p_k^\star}{96\sigma^2}T\lambda_k^2)}\right) \\
&\geqslant p_k^\star T\left(1 - \mathbb{P}\left(N_k^2 \leqslant \frac{p_k^\star}{12}T\right) - \frac{p_k^\star}{2}e^{-\frac{p_k^\star}{96\sigma^2}T\lambda_k^2)}\right) .
\end{aligned}$$

We finally upper bound $\mathbb{P}\left(N_k^2 \leqslant \frac{p_k^\star}{12}T\right) \leqslant \mathbb{P}(\bar{\mathcal{G}}_1)$, that we can again upper bound thanks to (4.19). We finally obtain that $B_T \sim p_k^\star T$ when $T \to +\infty$. Combining all the results developed in this part, we finally obtain the second statement of the theorem: when $\rho_\lambda > 0$, `P-SGOC` suffers no constraint violation asymptotically. □

4.7 Additional Experiments

In this section, we provide additional experiments to further support the results presented in Section 4.4.1. We first consider different value for the feasibility gap ρ_λ on a synthetic experiment with fixed distributions. Then, we redraw the Figure 4.1 presented in the main text but look only at the performance of `SGOC` to illustrate its behaviour with respect to the horizon T with better resolution. Lastly, we plot the results we obtain with the approach of [SSF22]. All code is written in python. Computations were run on a cluster with 10 cpus and 100 GB of RAM.

Impact of the feasability gap ρ_λ

In Figure 4.2, we study the impact of the feasibility gap ρ_λ on the performance of `SGOC`, `DOC`, `SPOC`, `BanditQ` in terms of violation $\mathcal{V}_T$ and excess regret $\mathcal{R}_T$ and the performance of `SGOC`, `DOC`, `SPOC`, `BanditQ`, `SGOC`, `P-SGOC` in terms of long term violation $\mathcal{V}_T$ and long term excess regret $\mathcal{R}_T^{LT}$. We take $K = 3$ arms, $\boldsymbol{\mu} = (0.8, 0.9, 0.7)$ with the feasibility ρ_λ varying in $\{0, 0.1, 0.5, 0.9\}$ and the horizon varying in $[10^2, 10^5]$. We set $\boldsymbol{\lambda} = \boldsymbol{\mu}(1 - \rho_\lambda)/K$. We take 200 seeds and report the mean value across seeds. Error bars represent the first and last decile.

When $\rho_\lambda = 0$, `SPOC`, `DOC` and `SGOC` behave like `DOC` and therefore have low regret but a constraint violation in $\sqrt{T}$. Looking at long term metrics, we see that all algorithms have low regret and positive fairness violation. BanditQ seems to get the best trade-off in this setting.

As the feasibility increases, `SGOC`, `DOC`, `SPOC`, and `SGOC` behave similarly as in the experiment in Figure 4.1 in the main text. In particular, we observe when $\rho_\lambda = 0.1$, the transition from optimism to pessimism of `SPOC` yielding a $\sqrt{T}$ regret but constant fairness violation. `SGOC` still has excellent performance with respect to long term metrics. The difference is the behavior of `BanditQ`. When $\rho_\lambda > 0$, `BanditQ` seems to have high long term excess regret but low long term violation. This is the opposite of what was observed in the experiment in Figure 4.1 in the main text. Such behaviour contrasts with the predictability of approaches like `SPOC` or `DOC`.

The regret and violation in $\sqrt{T}$ of `SGOC`

We redraw in Figure 4.3 the plot in Figure 4.1 but displaying only the performance of `SGOC`. We observe that the regret and constraints violation of `SGOC` evolve as $\sqrt{T}$ as expected from the analysis.

Results with LagrangeBwK [SSF22]

In Figure 4.4 we consider the same simulation setup as the one in Figure 4.1, including the LagrangeBwK algorithm from [SSF22]. We observe a linear constraint violation, showing that this implementation of the algorithm does not seem to converge to an optimal allocation for the problem and time horizon considered. We used the learning rates suggested in [SSF22], and for completeness we provide the implementation that we used in the supplementary material.

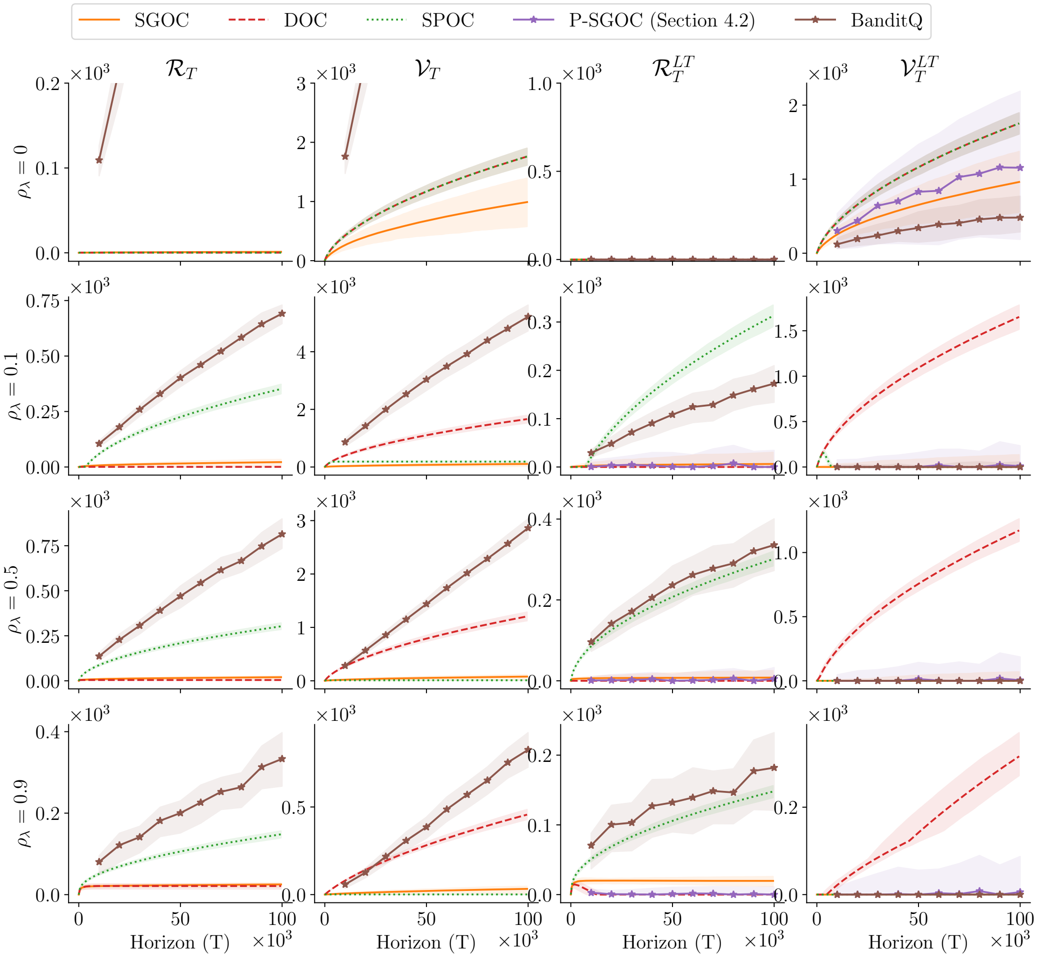


Fig. 4.2: Simulations with increasing feasibility gap ρ_λ. The plots in each row uses a different value of ρ_λ, the plots in each column represents a different metric.

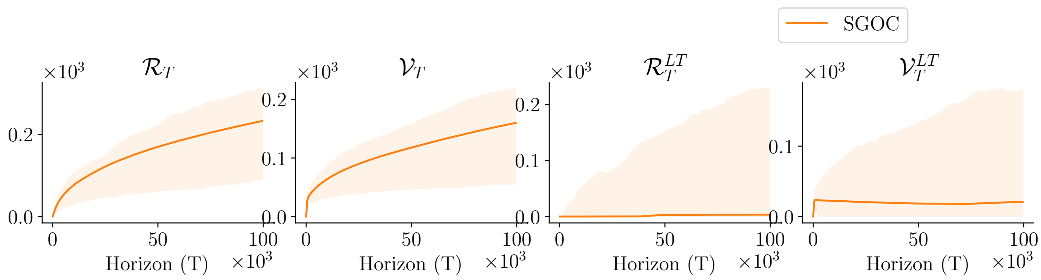


Fig. 4.3: Reproducing the simulation setup from Sinha [Sin23] focusing only on SGOC for better resolution

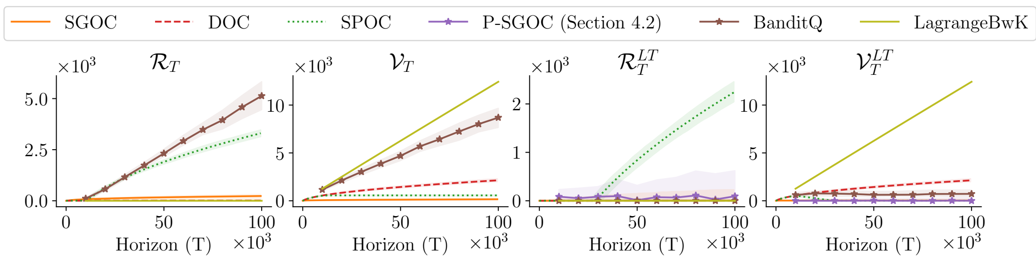


Fig. 4.4: Performance of LagrangeBwK [SSF22] on the simulation setup from Sinha [Sin23]

5

Prophet inequalities: competing with the top-ℓ is easy

Abstract. We explore a prophet inequality problem, where the values of a sequence of items are drawn i.i.d. from some distribution, and an online decision maker must select one item irrevocably. We establish that CR_ℓ the worst-case competitive ratio between the expected optimal performance of an online decision maker compared to that of a prophet who uses the average of the top ℓ items is exactly the solution to an integral equation. This quantity CR_ℓ is larger than $1 - e^{-\ell}$. This implies that the bound converges exponentially fast to 1 as ℓ grows. In particular for $\ell = 2$, $\mathrm{CR}_2 \approx 0.966$ which is much closer to 1 than the classical bound of 0.745 for $\ell = 1$. Additionally, we prove asymptotic lower bounds for the competitive ratio of a more general scenario, where the decision maker is permitted to select k items. This subsumes the k multi-unit i.i.d. prophet problem and provides the current best asymptotic guarantees, as well as enables broader understanding in the more general framework. Finally, we prove a tight asymptotic competitive ratio when only static threshold policies are allowed.

Contents

5.1 Introduction

Decision makers are frequently confronted with the arduous task of making crucial decisions with limited information. For example, when a seller wants to sell a limited number of items to a stream of customers, potential future customers with a high willingness to pay must be taken into account. How long should the seller wait before finally lowering its expectations, and to what price? Is the current customer likely to be the best we can hope to interact with? This common challenge is at the heart of many online selection problems [BE98]. Two of the simplest and most famous versions of this online selection problem are specific instances of optimal stopping problems [Cho+71]: the secretary problem where inputs are adversarial [Fer89], and prophet inequalities where inputs are random [Cor+19a]. In this work, we focus on the prophet inequality problem.

A classical way of measuring the performance of an online decision problem is to consider the so-called competitive ratio, the worst-case ratio between the performance of an online algorithm and the performance of an oracle that usually has access to more information than the decision maker. This has been the focus of many works, in online matching [Meh13], scheduling [MPT94], or metrical task systems [Bub+18] to name but a few, where explicit upper and lower bounds on the competitive ratio were provided. This metric makes it possible to design robust algorithms, that are able to always perform approximately well in any circumstances.

The classical prophet inequality dates back to the $1970s$, with [KS77] famously showing that a gambler allowed to select a single item using an optimal online algorithm can always recover at least half of the item value chosen by an omniscient prophet able to see the future, this $1/2$ factor being the best possible. Following this, [Sam84] proved that a simple threshold algorithm can actually achieve this competitive ratio of $1/2$. A refined prophet inequality was proved in [Hil83], who showed that if V is the maximum expected performance of an online algorithm, then the expected maximum value is always smaller than $2V - V^2$.

Following these works, variations with different assumptions on the item distributions were considered. If there are no assumptions on the joint distribution of the sequence of values, then [HK83] showed that the worst-case comparison between the gambler and the prophet can be arbitrarily bad in the number of items. Conversely, assumptions on the joint distribution can be strengthened: [HK82] are the first to consider the i.i.d. setting, in which the values of all items are independently drawn from the same distribution. An implicit upper bound of approximately 0.745 on the competitive ratio was proposed by [Ker86a] who reduces the computation of the worst-case competitive ratio to a finite dimensional optimization problem. This upper bound was proven to be tight by [Cor+17] through the construction of an explicit adaptive quantile algorithm that achieves this bound, showing that the worst-case competitive ratio in the i.i.d. setting is exactly the solution to an integral equation (with a numerical value of around 0.745). Some other works have further investigated the i.i.d. case: [PST22] show that this optimal competitive ratio can be approached with fewer different thresholds, and [JMZ22] show that the competitive ratio for a finite number of items can be computed as the solution to a linear program.

One important observation is that the worst-case instances tend to involve distributions that depend on the number of items and have a particularly heavy tail. This worst-case distribution does not correspond to the most commonly encountered distributions. In particular, for most distributions, the optimal online algorithm tends to perform better than what the worst-case instance suggests. As a result, some authors propose to use a different benchmark. [Ken85] and [Ker86b], for instance, studied the competitive ratio when the comparison is made with respect to a weaker prophet that receives the average reward of the top ℓ items. They prove that, in the case where valuations are independent but not necessarily identically distributed, the competitive ratio of any online algorithm cannot be larger than $1 - 1/(\ell + 1)$, and that this bound is attained. In this work, we consider the same benchmark of [Ken85; Ker86b], but with i.i.d. valuations as in [HK82; Ker86a; Cor+17]. We prove a lower bound on the competitive ratio and provide a quantile algorithm to solve the problem.

Contributions We consider a setting with $n \in \mathbf{b}N$ items whose valuations are $(X_1, \dots, X_n)$. The variables $(X_i)_{i\geq 1}$ are i.i.d. non-negative random variables with finite expectation drawn according to some distribution F, and we denote by $X_{(1)} \geqslant \dots \geqslant X_{(n)}$ their order statistics. The finite expectation assumption implies that the expectations of all order statistics are also finite. We consider online algorithms that observe valuations sequentially, and upon seeing the valuation X_i of an item i irrevocably decide to select the item and receive the value, or move to the next item. As a result, a valid algorithm induces a stopping time τ that corresponds to the selected item, and the expected performance of the algorithm is $E[X_\tau]$. We will compare the performance of an algorithm to the average valuation of the top ℓ items, $\mathbb{E}_{X\sim F}[\ell^{-1}\sum_{i\in[\ell]} X_{(i)}]$. For n and ℓ, we define the competitive ratio as:

$$\mathrm{CR}_\ell(n) = \inf_F \frac{\sup_\tau \mathbb{E}_{X\sim F}[X_\tau]}{\mathbb{E}_{X\sim F}\left[\frac{1}{\ell}\sum_{i\in[\ell]} X_{(i)}\right]}. \tag{5.1}$$

This corresponds to the competitive ratio achievable by an algorithm that knows the distribution F. Note that we always have $\mathrm{CR}_\ell(n) \leqslant 1$ by taking a constant distribution $X_i = 1$. We also define $\mathrm{CR}_\ell := \inf_{n\geqslant \ell} \mathrm{CR}_\ell(n)$ as the worst-case competitive ratio over all possible number of items. To better illustrate the practical implications of the model, consider the following two concrete interpretations. This competitive ratio can correspond to competing against an *imperfect prophet* who makes mistakes with some probability: the prophet can identify the top-ℓ items, but instead of selecting the maximum she selects one of the top-ℓ items uniformly at random. This competitive ratio could also correspond to a multi-unit setting with a budget of ℓ items, where a *stronger agent* competes against a regular prophet: the agent is allowed to re-select the same item multiple times. At least one optimal policy ends up selecting ℓ times the same item. This is immediate from the linearity of expectation: $\sup_{\tau_1,\dots,\tau_\ell} \mathbb{E}[\sum_{i\in[\ell]} X_{\tau_i}] = \sum_{i\in[\ell]} \sup_{\tau_i} \mathbb{E}[X_{\tau_i}] = \ell \sup_\tau \mathbb{E}[X_\tau]$.

Our first result is the following exact characterization of the worst-case competitive ratio:

Theorem 5.1. *For all positive integers ℓ, we have that CR_ℓ is the unique solution in $[0,1)$ to the integral equation*

$$\frac{1}{\ell!}\int_0^\infty \frac{\nu^{\ell-1}}{e^\nu\left(\frac{1}{\mathrm{CR}_\ell}-1\right)+\sum_{i=0}^{\ell}\frac{\nu^i}{i!}}d\nu = 1. \tag{5.2}$$

Our analysis is based on a non-trivial generalization of the quantile algorithm presented in [Cor+17], where this algorithm provides a lower bound on the competitive ratio. We show that maximizing the parameters of the proposed quantile algorithm (Algorithm 10) is equivalent to solving a non-linear discrete boundary value problem. We then show that this discrete boundary value problem corresponds to a continuous boundary value problem in the limit where n goes to infinity. This limiting competitive ratio lower bound yields the integral equation (5.2). Finally, we prove that the competitive ratio for a finite n is lower bounded by its limit as n grows, hence the limit performance of the algorithm is a lower bound on CR_ℓ. We obtain the matching upper bound by adapting a worst-case instance from [Liu+21a] to our setting.

ℓ	1	2	3	4	5
CR_ℓ	0.745	0.966	0.997	0.9998	0.999993

Tab. 5.1: First digits of CR_ℓ.

There are no explicit solutions to the integral equation (5.2), but this equation can be easily solved numerically. We report the first values of the competitive ratio CR_ℓ in Table 5.1. In the special case $\ell = 1$, we recover the integral equation of [Ker86a; Cor+17] that yields the competitive ratio of 0.745 for the classical i.i.d. prophet. What is striking is that for $\ell \geq 2$, the competitive ratio increases extremely fast towards 1. In particular, for $\ell = 2$, the competitive ratio is larger than 0.966, which is much closer to 1 than 0.745. For any distribution F, one has:

$$\sup_{\tau \text{ stopping time}} \mathbb{E}[X_\tau] \geqslant 0.745\, \mathbb{E}[X_{(1)}] \qquad \text{(result of [Cor+17])}$$

$$\sup_{\tau \text{ stopping time}} \mathbb{E}[X_\tau] \geqslant 0.966\, \mathbb{E}[\frac{1}{2}(X_{(1)} + X_{(2)})] \qquad \text{(our bound).}$$

The main reason for this difference is that the worst-case instances for $\ell = 1$ rely on the first and second maximum being very different. In fact, such an instance is relatively easy when using the benchmark $(X_{(1)} + X_{(2)})/2$. We observe numerically that CR_ℓ grows exponentially fast to 1 (roughly of the order of $1 - 10^{-\ell}$). Our second main result is to show that indeed the competitive ratio provably converges exponentially fast to 1 as ℓ grows:

Theorem 5.2. *For all positive integers ℓ, we have:*

$$\mathrm{CR}_\ell > 1 - e^{-\ell}. \tag{5.3}$$

Note that this bound is loose as can be seen with the values in Table 5.1 but still provides the exponential convergence rate to 1. Compared to the tight competitive ratio of $1 - 1/(\ell + 1)$ in [Ken85] without the i.i.d. assumption, the convergence towards 1 is noticeably faster. This exponential convergence rate explains why the jump from $\ell = 1$ to $\ell = 2$ was so marked.

All of our results are obtained through a quantile algorithm which is known to have additional benefits, such as being usable in a learning setting [RWW20]. Moreover, the algorithm only needs point-wise access to the quantile function, compared to the backward dynamic programming method that needs to compute an integral for each threshold.

There are multiple variants of the original setting of Theorem 5.1 that can be studied. A first possible extension is to consider a prophet that does not use the average of the top ℓ items, but instead uses a convex combination of the top ℓ items. Another possible variant is to compute the worst-case competitive ratio for the prophet. We prove some results for both of these extensions.

We also show that our algorithm can be extended with provable guarantees to a more general setting introduced by [Ken87] for the non i.i.d. case, where the decision maker is allowed to select k items. This setting encompasses the k multi-unit [Ala11; JMZ22] i.i.d. prophet problem when $\ell = k$. An algorithm that sequentially selects k items will induce a sequence of stopping times $(\tau_i)_{i\in[k]}$, for which if $\tau_i < \infty$ then $\tau_i < \tau_{i+1}$. We thus consider the following competitive ratio:

$$\mathrm{CR}_{k,\ell}(n) = \inf_F \frac{\sup_{\tau_1<\cdots<\tau_k} \mathbb{E}_{X\sim F}\left[\frac{1}{k}\sum_{i\in[k]} X_{\tau_i}\right]}{\mathbb{E}_{X\sim F}\left[\frac{1}{\ell}\sum_{i\in[\ell]} X_{(i)}\right]}. \tag{5.4}$$

We prove asymptotic guarantees on $\mathrm{CR}_{k,\ell}(n)$.

Theorem 5.3. *For all positive integers k and ℓ, we have:*

$$\liminf_{n\to\infty} \mathrm{CR}_{k,\ell}(n) \geqslant \frac{\mathrm{CR}_\ell}{k} \sum_{j\in[k]} \frac{1}{\prod_{t\in[j]} \theta_{j,\ell}}, \tag{5.5}$$

where $\theta_{1,\ell} = 1$ and $\theta_{2,\ell}, \ldots, \theta_{k,\ell}$ are the unique parameters such that the following boundary value problem admits a solution,

$$\begin{aligned} \frac{db^1(t)}{dt} &= \frac{\ell}{\mathrm{CR}_\ell} - \ell \cdot \gamma_{\ell+1} \circ \gamma_\ell(b^j(t)), \\ \frac{db^j(t)}{dt} &= \ell\left(\theta_{j,\ell} \cdot \gamma_{\ell+1} \circ \gamma_\ell^{-1}(b^{j-1}(t)) - \gamma_{\ell+1} \circ \gamma_\ell(b^j(t))\right), && \textit{for } 2 \leqslant j \leqslant k \\ b^j(0) &= 0, \quad b^j(1) = 1, && \textit{for } 1 \leqslant j \leqslant k, \end{aligned} \tag{5.6}$$

where $\gamma_x(t)$ and $\gamma_x^{-1}(t)$ are respectively the cumulative distribution function and the quantile function from a $\mathrm{Gamma}(x,1)$ random variable evaluated at t.

Remark that the first differential equation admits a solution if and only if CR_ℓ is the unique solution to the integral equation of Theorem 5.1.

Finally, we extend the analysis to the setting where the decision maker is restricted to use static thresholds policies. A static threshold policy $T \in \mathbf{b}R$ accepts items whenever $X_i \geqslant T$ and less than k items were already selected. More specifically, a static threshold T induces stopping times $\tau_1(T) = \inf\{i \in \mathbf{b}N \mid X_i \geqslant T\}$ and $\tau_i(T) = \inf\{i \in \mathbf{b}N \mid X_i \geqslant T, i > \tau_{i-1}\}$ for $i \geqslant 2$. When the decision maker and the prophet must select k and ℓ items respectively and the decision maker is restricted to static threshold policies we define the competitive ratio as

$$\mathrm{CR}^S_{k,\ell}(n) = \inf_F \frac{\sup_{T \in \mathbf{b}R} \mathbb{E}_{X\sim F}\left[\frac{1}{k}\sum_{i\in[k]} X_{\tau_i(T)}\right]}{\mathbb{E}_{X\sim F}\left[\frac{1}{\ell}\sum_{i\in[\ell]} X_{(i)}\right]}. \tag{5.7}$$

Using intermediary results derived for Theorems 5.1 and 5.3 and extending some of the analysis from [Cor+19a] and [AM21], we nearly characterize the exact competitive ratio for static thresholds:

Theorem 5.4. *For all positive integers k and ℓ, and $n \geqslant \max(k, \ell)$, we have:*

$$\left|\mathrm{CR}^S_{k,\ell}(n) - \frac{\sum_{j=1}^k \mathbb{P}(\mathrm{Gamma}(j,1) \leqslant \ell)}{k}\right| \leqslant O\left(\frac{1}{n}\right). \tag{5.8}$$

This result recovers a special case of the tight static threshold competitive ratio provided in [AM21] when $n \to \infty$ and with i.i.d. valuations, but is on other aspects more general by allowing for $\ell \neq k$.

Roadmap The rest of the paper is organized as follows. In Section 5.3, we present the quantile algorithm and the analysis of the competitive ratio for finite n. In Section 5.4 we show how to use the limit performance guarantees (as the number of items n goes to infinity) to construct the ODE and derive Theorem 5.1 and Theorem 5.2. We prove the matching upper bound of ℓ/c_ℓ on the competitive ratio in Section 5.5. In Section 5.6 we show how to extend our algorithm to the selection of k items and construct a corresponding system of k ODE to obtain Theorem 5.3. Section 5.7 deals with the static threshold setting and proves Theorem 5.4. Section 5.8 deals with some direct extensions to the setting considered, namely non uniform distributions for the *imperfect prophet*, and worst-case distribution against a prophet. The detailed proofs of all results are presented in Section 5.9. Finally, some additional related works are presented in Section 5.10.

5.2 Connection to fairness in allocation problems

While the present work treats a pure prophet inequality problem, the original motivation for the questions tackled in this article came from fairness considerations. This subsection is not present in the original publication and is only added to this thesis.

Now, to each item (or individual) $i \in [n]$, we associate for $p \in [0,1]$ the $i.i.d.$ random variable $A_i \sim \mathrm{Ber}(p)$, which corresponds to a sensitive attribute. For instance, $A_i = 0$ could correspond to a young person, and $A_i = 1$ to an older individual.

Say a company desires to hire 2 applicants among n agents, where hiring an employee corresponds to selecting the item (X_i, A_i) and receiving value X_i. We further require that the hiring has to be fair with respect to A_i, which we encode for some $\lambda \geqslant 0$ as receiving a penalty $-\lambda \cdot \mathbf{1}[(A_{\tau_1} = A_{\tau_2}) \cap (\max(\tau_1, \tau_2) < \infty)]$, which can be interpreted as having to pay a fine $-\lambda$ is two agents are selected and they have the same sensitive attribute (if only one is selected, we do not penalize the reward).

We now define the following competitive ratio with respect to λ, when the penalty is known to the decision maker:

$$\mathrm{CR}(\lambda) := \inf_{n,F,p} \sup_{\tau_1,\tau_2} \frac{\mathbb{E}[X_{\tau_1} + X_{\tau_2} - \lambda \cdot \mathbf{1}[(A_{\tau_1} = A_{\tau_2}) \cap (\max(\tau_1,\tau_2) < \infty)]]}{\max_{i \neq j} \mathbb{E}[X_i + X_j - \lambda \cdot \mathbf{1}[(A_i = A_j) \cap (\max(i,j) < \infty)]]}. \tag{5.9}$$

A first remark is that the prophet objective is now harder to evaluate, as the penalty needs to be taken into account.

For $\lambda = 0$ and $\lambda = \infty$, we can actually reduce the value of $\mathrm{CR}(\lambda)$ to specific values of $\mathrm{CR}_{k,\ell}$ which are studied in this paper.

Proposition 5.5. *If λ is known to the online decision maker, we have $\mathrm{CR}(\infty) = \mathrm{CR}_{1,1}$, and $\mathrm{CR}(0) = \mathrm{CR}_{2,2}$.*

Proof. The equality $\mathrm{CR}(\infty) = \mathrm{CR}_{2,2,}$ is immediate from the fact that without the penalty λ, this is simply a 2-unit $i.i.d.$ prophet inequality problem.

For $\mathrm{CR}(\infty) = \mathrm{CR}_{1,1}$, this is not as straightforward as it seems. Indeed, ideally, we would like to just run the algorithm which achieves $\mathrm{CR}_{1,1}$ twice in parallel, one accepting items with $A_i = 1$, and the other accepting items with $A_i = 0$. While the values X_i are still drawn from F, the number of occurrences of $A_i = 0$ and $A_i = 1$ are unknown in advance. But the original algorithm from [Cor+17] requires knowledge of the number of items faced. For the algorithm selecting items with

$A_i = 1$, we artificially assign the value of 0 to items with $A_i = 0$, and vice versa. Instead of running both quantile algorithms with F, we use them with the mixture $pF + (1-p)\delta_0$ for the first algorithm and with the mixture $(1-p)F + p\delta_0$, which is still i.i.d.. Because items with value 0 will be ignored, the two algorithms do not conflict and will not select the same item. Hence, by the main Theorem of [Cor+17], each modified online algorithm guarantees $\mathrm{CR}_{1,1}$ of their respective expected group maximum. That is to say $\mathbb{E}[X_{\tau_1}] \geqslant \mathrm{CR}_{1,1} \mathbb{E}[\max_{i,A_i=1} X_i]$ and $\mathbb{E}[X_{\tau_2}] \geqslant \mathrm{CR}_{1,1} \mathbb{E}[\max_{i,A_i=0} X_i]$. Summing these inequalities, we immediately get $\mathrm{CR}(\infty) \geqslant \mathrm{CR}_{1,1}$. The usual upper bound still applies to each individual algorithm, hence we do have $\mathrm{CR}(\infty) \leqslant \mathrm{CR}_{1,1}$, and the inequality holds. □

Hence, $\mathrm{CR}(\lambda)$ can be seen as an interpolation between $\mathrm{CR}_{1,1}$ and $\mathrm{CR}_{2,2}$. But while $\mathrm{CR}_{1,1}$ is known to be approximately 0.745 from [Cor+17], the value of $\mathrm{CR}_{2,2}$ was unknown. Hence, before even studying $\mathrm{CR}(\lambda)$, we needed to study $\mathrm{CR}_{2,2}$, and in Theorem 5.3 we provide asymptotic lower bounds which are the state of the art, and which we conjecture tight.

What if λ the penalty is unknown, in that it is revealed only at the end of the process? Let us define by CR_U the worst-case competitive ratio, when λ is revealed at the end of the process.

Proposition 5.6. *We have that* $\mathrm{CR}_U = \mathrm{CR}_{1,2}/2$.

Proof. The penalized prophet objective is always smaller than the unpenalized one. For the online agent, let us select a single item using the quantile algorithm which yields the lower bound $\mathrm{CR}_{1,2}$. Because a single item is collected, no penalty is incurred regardless of the value of λ. Therefore

$$\frac{\mathrm{ALG}}{\mathrm{OPT}} \geqslant \frac{\mathrm{ALG}}{\mathbb{E}[X_{(1)} + X_{(2)}]} = \frac{\mathbb{E}[X_{\tau_1}]}{\mathbb{E}[X_{(1)} + X_{(2)}]} \geqslant \frac{\mathrm{CR}_{1,2} \frac{\mathbb{E}[X_{(1)}+X_{(2)}]}{2}}{\mathbb{E}[X_{(1)} + X_{(2)}]} = \frac{\mathrm{CR}_{1,2}}{2},$$

which yields the lower bound.

For the upper bound, consider the sequence of worst-case distributions F_n which are tight for $\mathrm{CR}_{1,2}$, and let $p = 0$. If λ were to be equal to ∞, because the online agent is only faced with agents such that $A_i = 0$, only a single item can be selected. As λ is unknown and it could be ∞, this means that only a single item can be selected. If $\lambda = 0$, then the prophet objective is exactly $\mathbb{E}[X_{(1)} + X_{(2)}]$. Hence, in the worst case, with $p = 0$, λ unknown reduces to the $\mathrm{CR}_{1,2}$ benchmark, and inherits its competitive ratio from the minimizing sequence of the F_n. □

In $Theorem$ 5.1, we completely characterize $\mathrm{CR}_{1,2}$ and show that it is approximately 0.966. Hence, by the above proposition, we show that $\mathrm{CR}_U \approx 0.966$. This highlights that while the results presented in this paper do not seem directly related to fairness considerations, they are actually deeply linked and are necessary as a stepping stone.

The characterization of $\mathrm{CR}(\lambda)$ for arbitrary λ is still a work in progress.

5.3 Competitive ratio guarantees for a given number of items

Because the optimal online algorithm depends in a complicated fashion on the distribution, the direct analysis of the competitive ratio is hard. In this section, we construct an explicit simpler algorithm, and we derive an analysis of the competitive ratio of this algorithm. This algorithm builds on quantiles of the Beta distribution.

For the remainder of the paper and for simplicity of notations, we will consider that F is absolutely continuous[1] with respect to the Lebesgue measure and admits a density $f \geqslant 0$. The function F is the cumulative distribution function $F(x) = \mathbb{P}(X \leqslant x)$, and we denote by F^{-1} its quantile function: for each $p \in [0,1]$, $F^{-1}(p) = \inf\{x \in \mathbf{b}R \mid p \leqslant F(x)\}$. Most quantities depend on n and this will be omitted, except punctually when this makes the understanding clearer and will be denoted as $x(n)$ for some quantity x.

5.3.1 The quantile algorithm

We define the quantile algorithm Algorithm 10, which takes as an input the known distribution F, and an increasing sequence $0 = \varepsilon_0 < \cdots < \varepsilon_n = 1$. For each item i, this algorithm samples a quantile q_i from a Beta$(\ell, n-\ell)$ distribution truncated to ε_{i-1} and ε_i and selects item i if and only if $F(X_i) \geq 1 - q_i$. We denote by ALG_n the algorithm's expected performance for a sequence of n items. This algorithm generalizes the quantile algorithm described in [Cor+17] for the special case $\ell = 1$, and no mentions of Beta distributions were made.

[1]This assumption is standard for this type of analysis and simplifies the exposition. The proof can be adapted to general F by adding randomization between ties when the distribution has atoms, as in [Cor+17].

Algorithm 10 Quantile algorithm for ℓ

Input: Partition $(\varepsilon_i)_{0\leqslant i\leqslant n}$ of $[0,1]$, distribution F of the X_i.
for $i \in [n]$: **do**
  Draw q_i from $\mathrm{Beta}(\ell, n-\ell)$ truncated between ε_{i-1} and ε_i
  if $X_i \geqslant F^{-1}(1-q_i)$: **then**
    Accept item i and stop
  end if
end for

Before showing a bound on the performance of the algorithm, we introduce some notations regarding the Beta distribution. We recall that the density of a $\mathrm{Beta}(\ell, n-\ell)$ random variable is equal to

$$\psi_{\ell,n-\ell}(x) = \frac{x^{\ell}(1-x)^{n-\ell}}{B(\ell, n-\ell)}, \tag{5.10}$$

where $B(\ell, n-\ell) := (\ell-1)!(n-\ell-1)!/(n-1)!$ is the normalization constant. As q_i is drawn from a distribution truncated between ε_{i-1} and ε_i, we denote the normalizing factor of this $\psi_{\ell,n-\ell}$ truncated distribution by $\alpha_i = \int_{\varepsilon_{i-1}}^{\varepsilon_i} \psi_{\ell,n-\ell}(x)dx$. Finally, we define $a_i := \alpha_i \mathbb{E}[(1-q_i)]$, where $\mathbb{E}[(1-q_i)]$ is the expected probability of not selecting item i when observing it.

Our first result provides a bound on the performance of Algorithm 10 (valid for any sequence of ε_is), as a function of the quantities α_i and a_i defined above.

Proposition 5.7. *For* $\mathrm{OPT}_{\ell,n} = \mathbb{E}[\sum_{i\in[\ell]} X_{(i)}]$ *and* $\rho_i = \alpha_i^{-1}\prod_{j=1}^{i-1} a_j/\alpha_j$, *we have the inequality*

$$\frac{\min_{i\in[n]}\rho_i}{n}\,\mathrm{OPT}_{\ell,n} \leq \mathrm{ALG}_n \leq \frac{\max_{i\in[n]}\rho_i}{n}\,\mathrm{OPT}_{\ell,n}\,. \tag{5.11}$$

Sketch of proof. The proof decomposes in two steps. First, we compute an expression of ALG_n: remarking that the quantiles q_i are independent, we can show that the performance of ALG_n is equal to $\sum_{i\in[n]} \rho_i \int_{\varepsilon_{i-1}}^{\varepsilon_i} \psi_{\ell,n-\ell}(q)dq$. Second, we derive an expression for $\mathrm{OPT}_{\ell,n}$ as an expectation of the function $R(Q)$, where $R(q)$ is the expected reward of accepting an item with threshold $F^{-1}(1-q)$ and Q is some random variable. In [Cor+17], the distribution of Q was shown to be the density $(n-1)(1-q)^{n-2}$ in the special case $\ell = 1$. We prove by using the density for a general order statistic, that the right density in the case $\ell > 1$ is $\psi_{\ell,n-\ell}$. Because $\mathrm{OPT}_{\ell,n} = n\mathbb{E}_{q\sim\psi_{\ell,n-\ell}}[R(q)]$, we can take the minimum or the maximum over the ρ_i to obtain (5.11). A full proof is provided in Section 5.9.1. □

To obtain bounds on the competitive ratio of the algorithm, we can divide the above inequality by $\mathrm{OPT}_{\ell,n}/\ell$.

5.3.2 Optimizing the parameters of the algorithm

Looking at Proposition 5.7, we see that the ρ_is are functions of ε_i. A lower bound on the performance of the algorithm is therefore obtained if we can find the sequence ε_i which maximizes $\min_{i\in[n]} \rho_i$. A natural choice of ε_i would be to find a sequence such that all the ρ_i are equal, this would lead to an algorithm whose performance is *exactly* $(\max_i \rho_i/n)\,\mathrm{OPT}_n$. This would entail a lower bound on $\mathrm{CR}_\ell(n)$. It is, however, not clear whether such a sequence of ε_i exists. Moreover we need this sequence of ε_is to be increasing (*i.e.*, $\varepsilon_{i-1} < \varepsilon_i$) for the algorithm to be well defined.

We will first see that finding the ε_i such that $\rho_i = \rho_{i+1}$ is equivalent to solving a discrete boundary value problem on a non-linear transformation of the ε_i by an incomplete beta function. We then use this transformed problem to prove the existence of such an ε_i.

We introduce the variables $b_i = \beta_{\ell,n-\ell}(\varepsilon_i)$ which is a nonlinear transformation of ε_i, where

$$\beta_{\ell,n-\ell}(x) = \frac{\int_0^x t^{\ell-1}(1-t)^{n-\ell-1}dt}{B(\ell,n-\ell)}. \tag{5.12}$$

is the cumulative distribution function of $\psi_{\ell,n-\ell}$, also called the regularized incomplete beta function.

Because $\beta_{\ell,n-\ell}$ is strictly monotone as a distribution function with an associated positive density, it has an inverse, and we can recover ε_i with $\beta^{-1}_{\ell,n-\ell}(b_i)$. We also define the discrete difference operator Δ, with $\Delta[b_i] = b_{i+1} - b_i$. It is the discrete analogue of the continuous differentiation operator. Similarly, we define $\Delta^2[b_i] = \Delta[\Delta[b_i]] = b_{i+2} - 2b_{i+1} - b_i$.

Lemma 5.8. *All the ρ_i are equal if and only if the following difference equation holds for all b_i:*

$$\Delta[b_i] = -\frac{\ell}{n}\beta_{\ell+1,n-\ell} \circ \beta^{-1}_{\ell,n-\ell}(b_i) + b_1. \tag{5.13}$$

Sketch of proof. The main idea is that we can actually express α_i and a_i in terms of $\beta_{\ell,n-\ell}$. Indeed $\alpha_i = \beta_{\ell,n-\ell} = \beta_{\ell,n-\ell}(\varepsilon_i) - \beta_{\ell,n-\ell}(\varepsilon_{i-1}) = b_i - b_{i-1}$, and after some computations it can be showed that $a_i = (n-\ell)(\beta_{\ell,n+1-\ell}(\varepsilon_i) - \beta_{\ell,n+1-\ell}(\varepsilon_{i-1}))/n$. Then using that there is an exact recurrence relationship between $\beta_{\ell,n+1-\ell}$ and $\beta_{\ell,n-\ell}$, as well as between $\beta_{\ell+1,n-\ell}$ and $\beta_{\ell,n-\ell}$, we can obtain a non-linear second order difference equation. Then, for every $i \geqslant 2$, it is sufficient to sum these difference equations for all $j \leqslant i$, to obtain a recurrence relation directly on the b_i. The full proof can be found in Section 5.9.2. □

Remark. Note that obtaining an explicit recurrence relation of ε_{i+1} with 'simple' functions of the previous ε_i is difficult: developing the integrals a_i and α_i in ε_i and ε_{i+1} yields an implicit polynomial equation in ε_{i+1} parameterized by ε_i. This would entail finding roots of a sequence of polynomials of degree $n - \ell + 1$. The approach that we use here is to obtain an explicit recurrence relation on a non-linear transformation of the ε_i (the b_i), and not on the ε_i themselves. This non-linear transformation has no inverse expressible with 'simple' functions exactly whenever $\ell \geqslant 2$; this explains why the task is considerably more difficult than for the case $\ell = 1$ studied in [Cor+17]. This difficulty is one of the core obstacles to an extension of the work of [Cor+17].

It is quite remarkable that the recurrence relation is 'almost' linear, in that if $\beta_{\ell+1,n-\ell}$ were to be replaced with $\beta_{\ell,n-\ell}$ we would have recovered the identity when composing with $\beta^{-1}_{\ell,n-\ell}$. We will see below that when n goes to infinity we recover the Gamma distribution in the limit. As an aside when ℓ also goes to infinity we can recover the Normal distribution.

By using Lemma 5.8, we can therefore focus on proving the existence of the correct constant b_1 which will imply the required condition $\rho_{i+1} = \rho_i$.

Proposition 5.9. *There exists an increasing sequence* $0 = \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_{n-1} < \varepsilon_n = 1$ *and* $c_\ell(n) \in [\ell, \ell+1]$ *such that all the* ρ_i *are equal to* $n/c_\ell(n)$.

Sketch of proof. Because $\beta_{\ell,n-\ell}$ is a bijection from $[0,1]$ to $[0,1]$, $\beta_{\ell,n-\ell}(0) = 0$ and $\beta_{\ell,n-\ell}(1) = 1$, finding the right partition ε_i partition is equivalent to finding the right b_i partition. Due to the continuity of the recurrence relation, b_n is a continuous function of b_1, and the intermediate value theorem proves the existence. We must ensure that there is a solution in $[\ell/n, (\ell+1)/n]$ to guarantee the monotonicity of the b_i and thus of the ε_i. The full proof is provided in Section 5.9.3. □

Proposition 5.9 shows that there exists a sequence of ε_is such that our algorithm is well-defined and satisfies that $n/\rho_i = c_\ell(n)$ for all i. In particular, such an algorithm has a competitive ratio of $\ell/c_\ell(n)$. This shows that for all n, $\mathrm{CR}_\ell(n) \geqslant \ell/c_\ell(n)$.

In the remainder of the paper, we improve this result in two directions. First, the quantity $c_\ell(n)$ depends on n. In Section 5.4, we show how to obtain a guarantee that does not depend on n, by studying the limit as n grows. Second, the fact that $c_\ell(n) \in [\ell, \ell+1]$ implies a competitive ratio of at least $\ell\rho_1/n = \ell/c_\ell(n) \in [\frac{\ell}{\ell+1}, 1]$. The bound on $\ell/(\ell+1)$ is the same as the result from [Ken85; Ker86b] for the non-i.i.d. case. We then show in Section 5.4.3 that, in the i.i.d. setting, CR_ℓ is actually exponentially close to 1 when ℓ is large.

5.4 Competitive ratio guarantees as n grows

Until now, we have proven how to obtain guarantees that depend on the number of items. In this section, we first show that the worst-case for the competitive ratio is for large n. Then, we use a limiting ODE to characterize the competitive ratio given by our quantile algorithm when n goes to infinity.

5.4.1 $\mathrm{CR}_\ell(n)$ is minimized for very large n

By using our analysis of the previous section, we cannot directly conclude that $\mathrm{CR}_\ell(n)$ is a monotone function of the number of items n, nor that the competitive ratio might be small for large n. It might be possible that the value of CR_ℓ is actually reached for some n^* such that $\mathrm{CR}_\ell(n^*) = \mathrm{CR}_\ell$ (we know that this is not the case for $\ell = 1$, see [Cor+17]). Our Lemma 5.10 generalizes Lemma 3.2 from [Liu+21a] (that deals with the special case $\ell = 1$) and shows that we can always transform an instance with n items (F, n) to an instance with $2n$ items $(\widetilde{F}, 2n)$ that is at least as difficult as the (F, n) instance.

Lemma 5.10. *For any $n \geqslant \ell$, let $(X_1, \ldots, X_n)$ i.i.d. distributed according to F, and $(Y_1, \ldots, Y_{2n})$ i.i.d distributed according to $\sqrt{F}$. We have for τ_X and τ_Y the optimal stopping for respectively the X_i and Y_i that*

$$\frac{\mathbb{E}[X_{\tau_X}]}{\mathrm{OPT}_{\ell,n}(F)} \geqslant \frac{\mathbb{E}[Y_{\tau_Y}]}{\mathrm{OPT}_{\ell,2n}(\sqrt{F})}.$$

Sketch of proof. The original Lemma proves this for $\ell = 1$. They remark that for each X_i, we can simulate a draw of Y_{2i} and Y_{2i+1} conditionally on their maximum being equal to X_i. The item X_i is then accepted whenever Y_{2i} or Y_{2i+1} is accepted by τ_Y. This proves that $\mathbb{E}[X_{\tau_X}]$ is at least $\mathbb{E}[X_{\tau_Y}]$. This property holds for $\ell > 1$ as we are still only allowed to select a single item. What changes, however, is OPT_ℓ. We show that $\mathbb{E}[Y_{(\ell)}] \geqslant \mathbb{E}[X_{(\ell)}]$. It is sufficient to prove that $Y_{(\ell)}$ stochastically dominates

$X_{(\ell)}$, which will imply the inequality for the expectations. This can be proved by looking at the difference of the respective cumulative distribution functions and looking at the monotonicity of the derivative of the difference. The full proof can be found in Section 5.9.4. $\square$

Note that this lemma implies that $\mathrm{CR}_\ell(2n) \leq \mathrm{CR}_\ell(n)$. In particular, this implies that

$$\mathrm{CR}_\ell = \inf_{n \in \mathbf{b}N} \mathrm{CR}_\ell(n) = \liminf_{n \in \mathbf{b}N} \mathrm{CR}_\ell(n). \tag{5.14}$$

This explains why, in the rest of the section, we focus on the limiting behavior of the quantile algorithm as n goes to infinity. Note that this does not imply that $\mathrm{CR}_\ell(n)$ decreases with n.

5.4.2 Limiting ODE as n goes to infinity

We can now focus on analyzing the limit (as $n \to \infty$) of our discrete boundary value problem described by the recurrence relation in Equation (5.13). Let us first recall this difference equation using that $b_1 = c_\ell(n)/n$:

$$\Delta[b_i] = \frac{-\ell\beta_{\ell+1,n-\ell} \circ \beta^{-1}_{\ell,n-\ell}(b_i) + c_\ell(n)}{n}. \tag{5.15}$$

We show below that this difference equation converges to an ordinary differential equation as n goes to infinity, by using the property that the limit of a Beta random variable is a Gamma random variable. Recall that the cumulative distribution function of a $\mathrm{Gamma}(k,1)$ random variable is

$$\gamma_\ell(z) = \frac{\int_0^z t^{\ell-1} e^{-t} dt}{\Gamma(\ell)}, \tag{5.16}$$

where $\Gamma(\ell) = (\ell-1)!$ is the normalizing constant. This function is also called the regularized lower incomplete Gamma function. Because of the integral representation $\Gamma(\ell) = \int_0^\infty t^{\ell-1}e^{-t}dt$, it is clear that $\gamma_\ell(\infty) = 1$ and thus this is a proper distribution.

Lemma 5.11. *The sequence of functions $x \mapsto -\ell\beta_{\ell+1,n-\ell} \circ \beta^{-1}_{\ell,n-\ell}(x)$, defined on $[0,1]$, converges uniformly towards $-\ell\gamma_{\ell+1} \circ \gamma_\ell^{-1}(x)$ as n goes to infinity.*

Sketch of proof. We can prove the point-wise convergence by using the property that Beta random variables converge in distribution to Gamma random variables. Then to show that the convergence is uniform, we first show the monotonicity of the sequence of functions through usual formulas for beta functions. We can conclude by Dini's convergence theorem as the input space is compact. The full proof can be found in Section 5.9.5. □

Remark. The concise representation of the function of interest through Beta functions enables us to easily prove uniform convergence. Indeed, if we attempt to prove uniform convergence of first $\beta_{\ell,n-\ell}^{-1}$ and then $\beta_{\ell+1,n-\ell}$, this does not suffice as the output of the inverse of γ_ℓ is unbounded, so the input space of the last function is non-compact. Looking directly at the composition enables us to skip this difficulty, thus avoiding tedious technical computations, and piece-wise analysis.

Before stating formally the result, we start by giving the intuition on how to construct the limiting ODE. Lemma 5.11 suggests to approximate the difference equation (5.15) by

$$\frac{db(t)}{dt} = -\gamma_{\ell+1} \circ \gamma_\ell^{-1}(b(t)) + c_\ell, \tag{5.17}$$

where the solution b satisfies the boundary conditions $b(0) = 0$ and $b(1) = 1$, and where the constant c_ℓ is an unknown value that replaces $c_\ell(n)$.

As $c_\ell(n) \geqslant \ell$, we also consider that $c_\ell \geq \ell$. This implies that b is strictly increasing until at least the first t_1 for which $b(t_1) = 1$. So b is a bijection over $[0, t_1]$, and we can consider the inverse function $t(b)$ with $t(1) = t_1$ and $t(0) = 0$. Requiring that $t(1) = 1$ leads to the following integral equation:

$$\int_0^1 \frac{db}{-\ell\gamma_{\ell+1} \circ \gamma_\ell^{-1}(b) + c_\ell} = t(1) - t(0) = 1 - 0 = 1.$$

Using a change of variable $\nu = \gamma_\ell^{-1}$, we define c_ℓ as the constant which satisfies the following integral equation:

$$1 = \frac{1}{\Gamma(\ell)} \int_0^\infty \frac{\nu^{\ell-1}e^{-\nu}}{c_\ell - \ell\gamma_{\ell+1}(\nu)} d\nu = \frac{1}{\Gamma(\ell)} \int_0^\infty \frac{\nu^{\ell-1}}{e^\nu(c_\ell - \ell) + \ell\sum_{i=0}^{\ell} \frac{\nu^i}{i!}} d\nu. \tag{5.18}$$

Note that this last integral equation implies that $c_\ell > \ell$, as otherwise when ν goes to ∞ the integrand becomes equivalent to $\nu^{\ell-1}/\nu^\ell = 1/\nu$ which integrates to $\log(\nu)$ and diverges.

The next proposition formalizes the intuition, and shows the limit of $c_\ell(n)$ to indeed be the c_ℓ defined as the solution to this integral equation. The proof uses the same arguments as that in [Ker86a; JMZ22], with the main difference being the actual value of the limit, and proving the uniform convergence in Lemma 5.11, which has already been detailed. We defer the actual proof of the convergence to Section 5.9.6.

Proposition 5.12. *For c_ℓ the solution to Equation* (5.18)*, we have*

$$\lim_{n\to\infty} c_\ell(n) = c_\ell.$$

Combining this result with Equation (5.14) implies $\mathrm{CR}_\ell \geqslant \ell/c_\ell$. Remark that if c_ℓ is solution to Equation (5.18), then ℓ/c_ℓ is solution to Equation (5.2). This proves the first part of Theorem 5.1, namely that CR_ℓ is greater than the solution to the integral equation (5.2).

5.4.3 Asymptotic competitive ratio as ℓ grows

While the above integral equation does not lead to a close form expression for c_ℓ, it can be used to provide an easy characterization for the behavior of CR_ℓ as ℓ grows. Here, we recall and prove Theorem 5.2, which states that for all ℓ:

$$\mathrm{CR}_\ell > 1 - e^{-\ell}. \tag{5.19}$$

This result confirms what we observe in Table 5.1: the competitive ratio goes exponentially fast towards 1. To push the comparison deeper, we plot in Figure 5.1(a) the value of the $1 - \ell/c_\ell$ as a function of ℓ with a y-axis in log-scale. We observe that numerically $c_\ell \approx 1 - 10^{-\ell}$. This is closer to 1 than $1 - \exp(-\ell)$ predicted by Theorem 5.2, but still of the correct order. This does not tell us whether ℓ/c_ℓ is the best possible lower bound on CR_ℓ but shows that the competitive ratio must lie between $1 - \exp(-\ell)$ and 1 for all ℓ.

Proof of Theorem 5.2. For $c \geqslant \ell$, let us consider the integral in Equation (5.18):

$$\frac{1}{\Gamma(\ell)} \int_0^\infty \frac{\nu^{\ell-1}}{e^\nu(c-\ell) + \ell \sum_{i=0}^{\ell} \frac{\nu^i}{i!}} d\nu < \frac{1}{\ell\Gamma(\ell)} \int_0^\infty \frac{\nu^{\ell-1} e^{-\nu}}{(\frac{c}{\ell} - 1) + e^{-\nu} \sum_{i=0}^{\ell-1} \frac{\nu^i}{i!}} d\nu.$$

Let $h(\nu) := e^{-\nu} \sum_{i=0}^{\ell-1} \nu^i/i!$, then $h'(\nu) = -\nu^{\ell-1}e^{-\nu}/\Gamma(\ell)$. Hence

$$\begin{aligned}\frac{1}{\ell\Gamma(\ell)} \int_0^\infty \frac{\nu^{\ell-1} e^{-\nu}}{(\frac{c}{\ell} - 1) + e^{-\nu} \sum_{i=0}^{\ell-1} \frac{\nu^i}{i!}} d\nu &= \frac{-1}{\ell} \left[\log\left(\frac{c}{\ell} - 1 + h(\nu)\right)\right]_0^\infty \\ &= \frac{1}{\ell} \log\left(\frac{c}{c-\ell}\right).\end{aligned}$$

For $\tilde{c}_\ell$ such that $\log(\tilde{c}_\ell/(\tilde{c}_\ell-\ell))/\ell = 1$, then $\tilde{c}_\ell = \ell/(1-e^{-\ell})$. Moreover, because the integral in Equation (5.18) is decreasing in c, we have that $c_\ell < \tilde{c}_\ell$. Finally

$$\mathrm{CR}_\ell = \frac{\ell}{c_\ell} > \frac{\ell}{\tilde{c}_\ell} = 1 - e^{-\ell}. \qquad \square$$

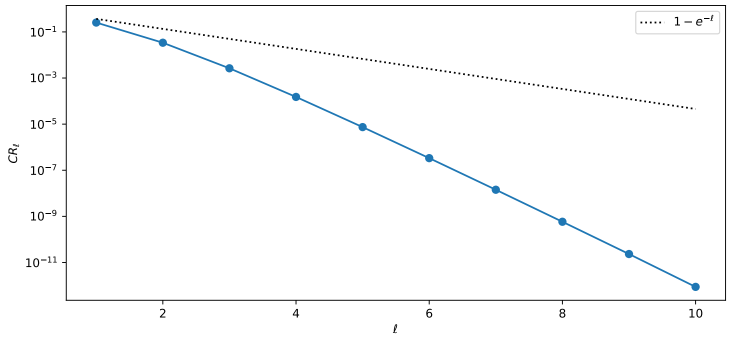


Fig. 5.1: CR_ℓ as a function of ℓ: the competitive ratio converges exponentially fast to 1.

5.5 Tightness of ℓ/c_ℓ

In this section, we show that the lower bound is actually tight, proving the result of Theorem 5.1. To do so, we adapt the worst-case instance for $\ell = 1$ from [Liu+21a] to an arbitrary ℓ.

In [Liu+21a], an equivalent random arrival model is used, where each item is assigned independently a uniform arrival time in $[0,1]$, and the decision maker observes the items in order of increasing arrival times. The optimal online policy $r_n(t)$ then corresponds to accepting an item with value X at time t if $r_n(t) > X$. The limit policy $\lim_{n\to\infty} r_n(t)$ is then studied, and the competitive ratio on an instance constructed from a differential equation is computed. Here, we similarly construct a worst-case example by leveraging the ODE Equation (5.17). For $b(t)$ the solution to the boundary value problem Equation (5.17), let $y(t) = 1 - b(t)$. The function $y(t)$ satisfies for all $t \in [0,1]$ the ODE

$$y'(t) = \ell\gamma_{\ell+1} \circ \gamma_\ell^{-1}(1 - y(t)) - c_\ell$$

with initial conditions $y(0) = 1$. Let $q \in (0,1)$ a parameter, $p = \exp(-\gamma_\ell^{-1}(1-y(q)))$, and

$$H = \frac{1}{y'(q)\log(p)} - \int_q^1 \frac{ds}{y'(s)}.$$

We define a threshold policy $r(t)$ as

$$r(t) = -\int_t^1 \frac{ds}{y'(s)} \quad \text{for } t \in [q,1] \text{ and } r(t) = H \text{ for } t \in [0,q).$$

Finally, we define F_q the distribution of the maximum as

$$F_q(x) = \begin{cases} \exp(-\gamma_\ell^{-1}(1 - y(r^{-1}(x))) & \text{for } t \in [0, r(q)] \\ p \quad \text{for } t \in (r(q), H] \\ 1 \quad \text{for } t \in (H, \infty). \end{cases} \tag{5.20}$$

This distribution is constructed such that r is optimal in the limit. The X_i are thus distributed according to $F_q^{1/n}$.

Proposition 5.13.

$$\lim_{q\to 0} \lim_{n\to\infty} \sup_{\tau \text{ stopping time}} \frac{\mathbb{E}_{X\sim F_q^{1/n}}[X_\tau]}{\mathbb{E}_{X\sim F_q^{1/n}}\left[\frac{1}{\ell}\sum_{i=1}^{\ell} X_{(i)}\right]} = \frac{\ell}{c_\ell}. \tag{5.21}$$

Sketch of proof. The first step is to show that the threshold policy $r(t)$ defined above is indeed the limit online optimal policy as $n \to \infty$, for this particular instance F_q. Then, we compute the value of the competitive ratio as q goes to 0, and show through integration by parts that this ratio equals ℓ/c_ℓ, using the characterization of c_ℓ with the integral equation (5.18). The full proof can be found in Section 5.9.8 □

The above property implies that $\mathrm{CR}_\ell \leqslant \ell/c_\ell$, which concludes the proof of Theorem 5.1.

5.6 Extension to the selection of k items

Until now we have assumed, that the number of items that can be selected by the decision maker is only 1. Here, we consider an extension of the problem where the decision maker sequentially selects k items, and where all the items can be selected at most once. This corresponds to the more general setting of [Ken87], which encloses the previous model.

5.6.1 General algorithm and guarantees

This section focuses on giving guarantees on $\mathrm{CR}_{k,\ell}(n)$ for $k > 1$. This time, the algorithm needs to be adaptive not only in n but also in k. We define the quantile algorithm Algorithm 11, which takes as an input the known distribution F, and for $j \in [k]$, the increasing sequences $0 =: \varepsilon^j_{j-1} < \varepsilon^j_j < \cdots < \varepsilon^j_n := 1$. The algorithm is the direct extension of Algorithm 10. The only difference is that, in this new version, the algorithm uses thresholds that depend both on the number of items already observed $(i-1)$ and on the number of items already selected $(j-1)$.

Algorithm 11 Quantile algorithm for (k, ℓ)

Input: Partition $(\varepsilon^j_i)_{j-1 \leqslant i \leqslant n}$ of $[0,1]$ for all $j \in [k]$, distribution F of the X_i.
$j \leftarrow 1$ *// We are currently selecting the jth item.*
for $i \in [n]$: **do**
  Draw q^j_i from $\mathrm{Beta}(\ell, n-\ell)$ truncated between ε^j_{i-1} and ε^j_i
  if $X_i \geqslant F^{-1}(1 - q^j_i)$ **then**
    Accept item i
    if $j = k$ **then**
      Stop *// because we selected k items*
    else
      $j \leftarrow j + 1$
    end if
  end if
end for

Similarly to earlier, we define $\alpha^j_i := \int_{\varepsilon^j_{i-1}}^{\varepsilon^j_i} \psi_{\ell,n-\ell}(x)dx = \beta_{\ell,n-\ell}(\varepsilon^j_i) - \beta_{\ell,n-\ell}(\varepsilon^j_{i-1})$, and $a^j_i = \alpha^j_i \mathbb{E}[(1-q^j_i)]$ where $\mathbb{E}[(1-q^j_i)]$ is the expected probability of not selecting item i when we are observing it and have currently already selected $j-1$ items. Similarly to Proposition 5.7, we can obtain a guarantee on $\mathrm{ALG}_{k,n} := \mathbb{E}[\sum_{j \in [k]} X_{\tau_j}]$, the performance of Algorithm 11, with respect to $\mathrm{OPT}_{\ell,n}$.

Proposition 5.14. *We have the inequality*

$$\frac{\sum_{j=1}^k \min_{i \in \{j,\dots,n\}} \rho^j_i}{n} \mathrm{OPT}_{\ell,n} \leqslant \mathrm{ALG}_{k,n} \leqslant \frac{\sum_{j=1}^k \max_{i \in \{j,\dots,n\}} \rho^j_i}{n} \mathrm{OPT}_{\ell,n}, \quad (5.22)$$

where

$$\rho^j_i = \frac{1}{\alpha^j_i} \sum_{1 \leqslant t_1 < \cdots < t_{j-1} \leqslant i-1} \left[\prod_{s=1}^{j-1} \left(1 - \frac{a^s_{t_s}}{\alpha^s_{t_s}}\right) \prod_{r_1=1}^{t_1-1} \frac{a^1_{r_1}}{\alpha^1_{r_1}} \cdots \prod_{r_j=t_{j-1}+1}^{i-1} \frac{a^j_{r_j}}{\alpha^j_{r_j}} \right]. \quad (5.23)$$

Sketch of proof. This is similar to Proposition 5.7, the main difference being that the performance of the algorithm conditionally on the q_i being already drawn is more difficult to express. See Section 5.9.9 for the proof. $\square$

Remark that the ρ_i^j are only defined for $i \geqslant j$, as time j is the first possible time for the j-th item to be selected. If for all $j \in [k]$, we have that the ρ_i^j are all equal in i, meaning that $\rho_i^j = \rho_j^j$, then this readily implies from the previous proposition that $\mathrm{ALG}_{k,n} = \frac{\sum_{j\in[k]} \rho_j^j}{n} \mathrm{OPT}_{\ell,n}$, and that

$$\mathrm{CR}_{k,\ell}(n) \geqslant \frac{\ell}{k} \frac{\sum_{j\in[k]} \rho_j^j}{n} \mathrm{OPT}_{\ell,n}, \tag{5.24}$$

For all $j \in [k]$, we want to find the ε_i^j that equalizes all the ρ_i^j across the different i. Here, looking at the expression of ρ_i^j, finding any meaningful recurrence relationship might seem hopeless. However, remark that the probability of reaching item i while waiting to select item j only depends on the ε_r^t for $t \leqslant j$, and thus so does ρ_i^j. This implies that in order to equalize the ρ_i^j across i, it must be done inductively over j: first select the ε_i^1 such that $\rho_i^j = \rho_1^j$, then select the ε_i^2 such that $\rho_i^2 = \rho_2^2$, and so on. We show that, this is equivalent to a system of k difference equations in a non-linear transformation of the ε_i^j.

Lemma 5.15. *For $b_i^j = \beta_{\ell,n-\ell}(\varepsilon_i^j)$, the condition that for all fixed j the ρ_i^j are equalized, is equivalent to the following system of difference equations over the b_i^j:*

$$\Delta[b_i^j] = -\frac{\ell}{n}\left(\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(b_i^j) - \frac{\rho_{j-1}^{j-1}}{\rho_j^j}\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(b_i^{j-1})\right) + b_j^j, \qquad \textit{for } j \geqslant 2,\ i \geqslant j \tag{5.25}$$

$$\Delta[b_i^1] = -\frac{\ell}{n}\left(\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(b_i^1)\right) + b_1^1, \qquad \textit{for } i \geqslant 1. \tag{5.26}$$

Sketch of proof. The most difficult part, is to actually identify the recurrence relationship between the α_i^j and a_i^j that the equality of the ρ_i^j imposes. While it was immediate for $j = 1$ as we simply had $\alpha_{i+1}^1 = a_i^1$, it is not clear what relationship can be obtained for $j \geqslant 2$. Fortunately, a simple relation is obtained by incorporating for j the previous constant ρ_{j-1}^{j-1}. Indeed, this is equivalent for $j > 1$ to $\alpha_{i+1}^j = a_i^j + (\alpha_i^{j-1} - a_i^{j-1})(\rho_{j-1}^{j-1})/\rho_j^j$. From there, obtaining the recurrence relation is as before based on the properties of the Beta function. For the proof, see Section 5.9.10. $\square$

We define $\theta_{j,\ell}(n) := \rho_{j-1}^{j-1}/\rho_j^j$. From this recurrence relation, we can prove the existence of the solution to the system of boundary discrete value problem by using the exact same continuity argument as in Proposition 5.9. The proof is however much more technically involved, and requires using several estimates of b_j^j and $\theta_{j,\ell}(n)$ when n grows large.

Proposition 5.16. *There exists some $n_0 \in \mathbb{N}$, such that for $n \geqslant n_0$, there exist k increasing sequences $0 = \varepsilon_{j-1}^j < \varepsilon_j^j < \cdots < \varepsilon_n^j = 1$ for $j \in [k]$ and $c_{j,\ell}(n)$ such that: for a given j all the ρ_i^j are equal, $b_j^j = c_{j,\ell}(n) \cdot n^{-\ell \cdot (((\ell+1)/\ell)^j - 1)}$ with $c_{j,\ell}(n)$ being bounded between two positive constants independent of n.*

Sketch of proof. The quantities b_n^j are still the composition of continuous functions (yet different functions for each $i \geqslant j+1$), so the same intermediate value argument can be applied to prove existence. If $\theta_{j,\ell}(n)$ is too big or too small compared to some constants, so is b_n^j compared to 1. Regarding the exponent in n, it can be proved by induction using the relation between $c_{j,\ell}(n)$ and $\theta_{j,\ell}(n)$. The monotonicity comes from the requirement that all the ρ_i^j are equal and thus must be of the same sign. See section 5.9.11 for the full proof. $\square$

This proposition suggests that the discrete boundary value problem can be approximated in the limit by the continuous boundary value problem in Equation (5.6). By Proposition 5.12 and Theorem 5.1 we already have that $\lim_{n\to\infty} c_{1,\ell}(n) = \ell/\operatorname{CR}_\ell$.

The goal is then to find $(\theta_{2,\ell}, \ldots, \theta_{k,\ell})$ such that this non-linear ODE system admits a solution $b = (b^1, \ldots, b^k)$ over $[0,1]$, which will be unique. Note that these constants can be found by sequentially solving the j-th ODE and finding the j-th relevant constant.

Through Proposition 5.14, solving the discrete boundary value problem for a finite n directly translates to a lower bound on the competitive ratio $\operatorname{CR}_{k,\ell}(n)$. To show that the limiting competitive ratio can also be lower bounded, we must show that the solutions $\theta_j(n)$ to the discrete problem converge, which naturally ends up being the solution to the above continuous boundary value problem.

Proposition 5.17. *There exists unique $\theta_{2,\ell}, \ldots, \theta_{k,\ell}$ constants such that the boundary value problem in Equation* (5.6) *admits a solution. We also have that $\lim_{n\to\infty} \theta_{j,\ell}(n) = \theta_{j,\ell}$, and $\theta_{j,\ell} \geqslant 1$. Moreover, this also implies the convergence of $c_{j,\ell}(n)$ toward a constant $c_{j,\ell}$, and for all $j \geqslant 2$ the relationship:*

$$\theta_{j,\ell} = \frac{(\ell+1)c_{j,\ell}}{\ell(\ell!)^{1/\ell} \cdot c_{j-1,\ell}^{1+1/\ell}}. \tag{5.27}$$

Sketch of proof. The main idea is to couple the convergence of the Euler scheme with the uniform convergence of the drift function, as the difference equations are Euler schemes that use $\beta_{\ell,n-\ell}$ instead of γ_ℓ. The sequence $\theta_{j,\ell}(n)$ is bounded, and thus admits at least one accumulation point. If we prove the convergence, on a subsequence if needed, of the discrete solution towards the ODE with θ an accumulation point of $\theta_{j,\ell}(n)$ then this shows the existence of a constant such that the continuous boundary value problem admits a solution. This constant is then shown to be unique, which proves that there is only a single accumulation point, and the sequence must converge. However, a key technical difficulty is that the Euler method requires Lipschitzness of the drift function, which is not true for $\gamma_{\ell+1} \circ \gamma_\ell^{-1}$ over $[0,1]$. Thus we must use refined arguments proving the convergence on $[0,1-\varepsilon]$ to then extend the convergence over $[0,1]$. The full proof can be found in Section 5.9.12. $\square$

We now combine the above results to prove Theorem 5.3. Letting $\theta_{1,\ell} = 1$ for ease of notation, we have that $\rho_i^j = \rho_j^j = 1/(c_{1,\ell} \cdot \prod_{t\in[k]} \theta_{t,\ell})$, and therefore applying $\liminf$ on the inequality from Proposition 5.14 and using that the $\theta_{j,\ell}$ converge from Proposition 5.17,

$$\liminf_{n\to\infty} \mathrm{CR}_{k,\ell}(n) \geqslant \liminf_{n\to\infty} \frac{\sum_{j\in[k]} \rho_j^j(n)}{n} = \lim_{n\to\infty} \frac{\ell}{k} \frac{1}{c_{1,\ell}(n)} \sum_{j\in[k]} \frac{1}{\prod_{t\in[j]} \theta_{t,\ell}(n)} = \frac{\mathrm{CR}_\ell}{k} \sum_{j\in[k]} \frac{1}{\prod_{t\in[j]} \theta_{t,\ell}}, \tag{5.28}$$

which concludes the proof.

5.6.2 Numerical results for general setting

Using numerical optimization, we compute the constants $\theta_{j,\ell}$ and provide in Table 5.2 the numerical value of the asymptotic lower bound on $\mathrm{CR}_{k,\ell}(n)$ from Theorem 5.3. We also display the solution to the continuous boundary value problem for $k = \ell = 5$ in Figure 5.2.

We do not have a proof of the asymptotic tightness of Equation (5.28), nor do we have a proof that $\mathrm{CR}_{k,\ell} = \liminf_{n\to\infty} \mathrm{CR}_{k,\ell}(n)$, although we conjecture that both statements are true. The asymptotic tightness is harder to show due to the lack of simple integral characterization, and the same strategy as Lemma 5.10 cannot be used for a selection budget greater than 1. Nevertheless, we still show that the infinite-dimensional optimization problem of computing $\mathrm{CR}_{k,\ell}(n)$ is simpler than it appears, and can be reduced to a finite-dimensional optimization by applying the

$k \backslash \ell$	1	2	3	4	5
1	0.745	0.966	0.997	0.9998	0.999993
2	0.486	0.829	0.964	0.995	0.9995
3	0.332	0.645	0.864	0.964	0.993
4	0.24997	0.498	0.724	0.885	0.964
5	0.19997	0.3998	0.596	0.772	0.898

Tab. 5.2: First digits of $(\mathrm{CR}_\ell/k) \sum_{j \in [k]} \prod_{t \in [j]} \theta_t^{-1}$.

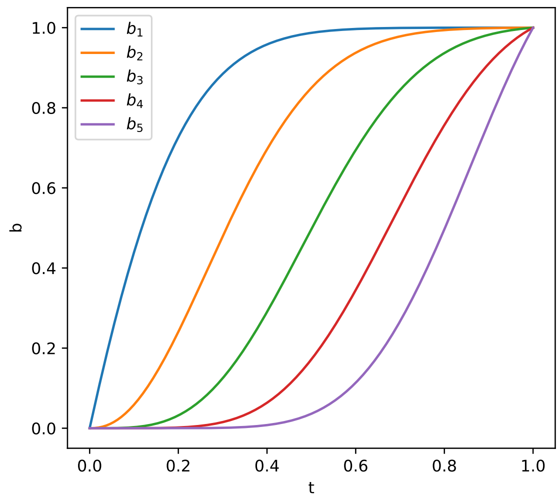


Fig. 5.2: Solution to the continuous boundary value problem for $k = \ell = 5$.

balayage technique from [HK82]. Solving this optimization problem numerically would then provide valid upper bounds on $\mathrm{CR}_{k,\ell}(n)$.

Proposition 5.18. *The value of* $\mathrm{CR}_{k,\ell}(n)$ *is attained by a discrete distribution with a support of* $2 + k(k-1)/2 + k(n-k)$ *points on* $[0,1]$.

See section 5.9.15 for the proof. This proposition is actually stronger than a similar result of [JMZ22] (Lemma 7.2), who show that for the (k,k) case, using an increasingly finer discretization over values to solve the optimization problem $\mathrm{CR}_{k,k}(n)$ approximates well the optimal value. Here we have shown that not only can this be extended to any general (k,ℓ) setting, but mainly that at least one minimum distribution *must* lie in a discretization linear in n, and it is unnecessary to make the discretization any finer.

5.7 Static thresholds

We now restrict the competitive ratio analysis to the set of static threshold policies. Similarly to previous works [AM21], we allow for random tie breaks when the distributions are discrete. For simplicity, the exposition will use continuous distribution.

In the i.i.d. single item setting, it has been known that the threshold $F^{-1}(1-1/n)$ achieves a competitive ratio of $1-1/e$. A simple alternate proof of this fact was presented in [Cor+19a] using the representation of $\mathbb{E}[\max_i X_i]$ as the expectation of $nR(Q)$ for Q distributed according to some distribution, and the Jensen inequality. As we have generalized this result and obtained that $\mathrm{OPT}_{\ell,n}$ is equal to $n\mathbb{E}[R(Q)]$ with Q distributed according to $\mathrm{Beta}(\ell, n-\ell)$, we use the same method to prove the following lower bound:

Proposition 5.19. *The performance of the algorithm that uses static threshold $T = F^{-1}(1-\ell/n)$ is greater than*

$$\frac{\sum_{j=1}^{k}\gamma_j(\ell)}{k} - \frac{1}{n}\left(1-\gamma_j(\ell) - \frac{\ell^{j-1}e^{-\ell}}{(j-1)!}\right) - o\left(\frac{1}{n^2}\right). \tag{5.29}$$

For the full proof see Section 5.9.13. Compared to the proof of $1-1/e$ in [Cor+17], multiple additional algebraic manipulations are necessary. This result is actually even more precise, in the sense that the expected reward of the j-th item is up to the error term exactly $\gamma_j(\ell)/\ell$. One aspect of this result that is remarkable, is that the threshold *only depends on* ℓ *and not on* k. This is quite surprising as this suggests for the decision maker to target the expected demand of the prophet in order to achieve a good competitive ratio.

To obtain an upper bound, we can adapt results from [AM21] which deals with the k multi-unit static threshold prophet secretary problem. It so happens that their worst-case instance is i.i.d.. We use the following modified example: Let F^* be the distribution such that $X = 1$ with probability $1-1/n^2$, and $X = nW_{k,\ell}$ with probability $1/n^2$ where

$$W_{k,\ell} = \frac{\ell^2}{k}\frac{\mathbb{P}(\mathrm{Poisson}(\ell) < k)}{\mathbb{P}(\mathrm{Poisson}(\ell) > k)}. \tag{5.30}$$

This example provides an asymptotically tight upper bound.

Proposition 5.20. *For* (T^*, p^*) *the optimal static threshold and random tie-break under* F^**, we have*

$$\lim_{n\to\infty} \frac{\mathbb{E}_{X\sim F}\left[\frac{1}{k}\sum_{i\in[k]} X_{\tau_i(T^*)}\right]}{\mathbb{E}_{X\sim F}\left[\frac{1}{\ell}\sum_{i\in[\ell]} X_{(i)}\right]} = \frac{\sum_{j=1}^{k}\gamma_j(\ell)}{k}. \tag{5.31}$$

The full proof can be found in Section 5.9.14. The combination of these two results immediately yields Theorem 5.4.

Remark. For the special case $k = 1$, the same arguments of Lemma 5.10 can be applied. For any threshold T, the expected performance of this threshold over an instance $X_1, \ldots, X_n \sim F$ will be greater than for an instance $Y_1, \ldots, Y_{2n} \sim \sqrt{F}$, as playing on the first instance with a static threshold is equivalent to playing against $(\max(Y_{2i}, Y_{2i+1}))_{i\in[n]}$. Therefore, $\mathrm{CR}_\ell^S := \inf_{n\geqslant \ell} \mathrm{CR}_\ell^S(n)$ and thus $\mathrm{CR}_\ell^S = \gamma_1(\ell) = 1 - e^{-\ell}$.

A direct consequence of the previous remark and Theorem 5.2, is that for $k = 1$ there is a positive gap between the worst-case performance of static threshold policies compared to dynamic threshold policies:

Corollary 5.21. *For* $\ell \in \mathbb{N}$, $\mathrm{CR}_\ell > \mathrm{CR}_\ell^S$.

We compute $\sum_{j=1}^{k}\gamma_j(\ell)/k$ for different k and ℓ and represent them in the left plot of Figure 5.3. We observe that when either k or ℓ or both grow large, the competitive ratio goes towards $\min(\ell/k, 1)$. In addition, the convergence to $\min(\ell/k, 1)$ seems to be the slowest for $k = \ell$ and otherwise exponential, away from $k = \ell$, as can be observed on the right plot of Figure 5.3.

This result is intuitive, and we present a simple explanation for $\ell = 1$ and k arbitrary. For the single item i.i.d. worst case instance, the first maximum is very far from the second maximum, and thus all other order statistics. For this specific distribution, while having a large k allows a greater probability of selecting the actual maximum, all the other selected values will be negligible compared to the maximum. Hence the expected reward of the decision maker will approach $\mathbb{E}[\max_i X_i]$, and the mean reward $\mathbb{E}[\max_i X_i]/k$ which leads to a competitive ratio of order $1/k$.

An interesting open question is whether similar guarantees extend to the prophet secretary setting, where distributions are not identical anymore and arrive in random order. In [AM21], the $k = \ell$ case is studied, which proves that expected demand

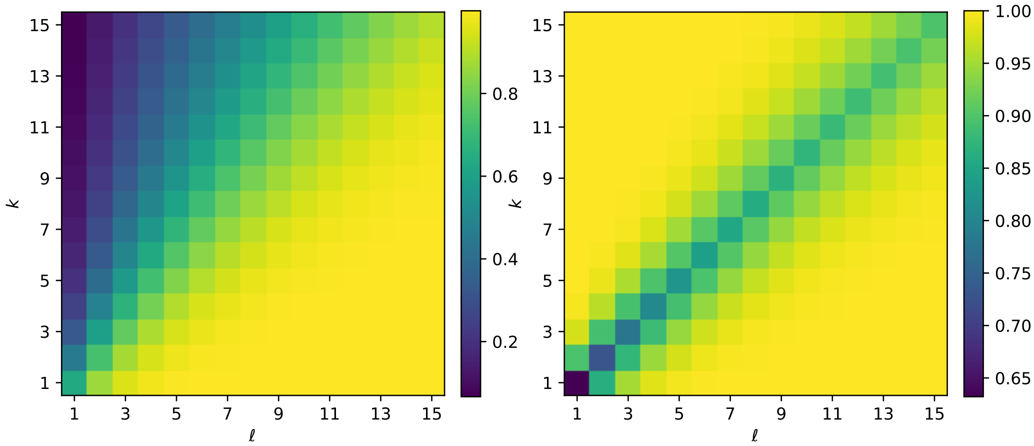


Fig. 5.3: For $k, \ell \in [15]$: on the left $\mathrm{CR}^S_{k,\ell}$, on the right $\mathrm{CR}^S_{k,\ell} \cdot \min(\ell/k, 1)$

policies are not tight for $k \leqslant 4$, but are tight whenever $k = \ell > 4$. The proof of the general $k \neq \ell$ setting presented here relies heavily on the i.i.d. assumption, but it is likely that the values of the competitive ratio in the prophet secretary setting remain similar to the i.i.d. setting

5.8 Extensions

5.8.1 Selecting order statistics with decreasing distributions

In this subsection we will only consider the case $k = 1$ for exposition purposes, the case $k > 1$ can be treated similarly.

Up until now, we have assumed as a benchmark $\mathbb{E}[\sum_{i \in [\ell]} X_{(i)}/\ell]$, which can be reformulated according to our *imperfect prophet* as $\mathbb{E}[X_{(S)}]$ where $S \sim \mathrm{Unif}([\ell])$. Can we say anything when the probability distribution of S is not uniform? We can generalize this benchmark to any decreasing probability mass function $\mathbb{P}(S = s) = p_s$ for $s \in \mathbf{b}N$, where $p_1 \geqslant p_2 \geqslant \ldots \geqslant p_\ell > p_{\ell+1} = 0$ and such that $\sum_{s \in [\ell]} p_s = 1$. Remark that this implies $p_1 \geqslant 1/\ell$. Let $\mathbf{p} = (p_1, \ldots, p_\ell, \ldots)$, and $\mathrm{CR}(\mathbf{p})$ be the competitive ratio with the benchmark $\mathbb{E}[X_{(S)}]$ where S is distributed according to $\mathbf{p}$ over $\mathbf{b}N$. Clearly, $\mathrm{CR}_\ell = \mathrm{CR}((\mathbf{1}_\ell, 0, \ldots)/\ell)$. For simplicity and to avoid convergence issues in n, we only allow for finite distributions, i.e. $\sup\{i \mid p_i > 0\} < \infty$, but the model could still be defined even for a countably infinite distribution by setting $X_{(s)} = 0$ if $s > n$.

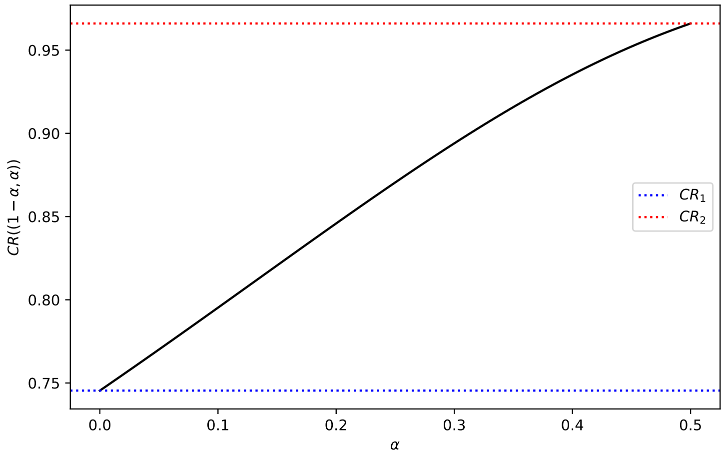


Fig. 5.4: Competitive ratio for $\mathbf{p} = (1-\alpha, \alpha)$, for $\alpha \in [0, 1/2]$.

Proposition 5.22. *For $\mathbf{p} = (p_1, \ldots, p_\ell, \ldots) \in \mathbf{b}R_+^{\mathbf{b}N}$, with $\mathbf{p}$ non-increasing, $\ell := \sup\{i \mid p_i > 0\} < \infty$ and $\sum_{i \in \mathbf{b}N} p_i = 1$, we have*

$$\mathrm{CR}(\mathbf{p}) \geqslant \frac{1}{p_1 \cdot c(\mathbf{p})}, \tag{5.32}$$

where $c(\mathbf{p})$ is the unique solution in $[1/p_1, \infty)$ to the integral equation

$$1 = \int_0^\infty \frac{e^{-\nu} \sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \frac{\nu^{s-1}}{(s-1)!}}{c(\mathbf{p}) - \sum_{s \in [\ell]} \frac{p_s - p_{s+1}}{p_1} \cdot s \cdot \gamma_{s+1}(\nu)} d\nu.$$

Proof. The core difference compared to CR_ℓ, is that $\psi_{\ell,n-\ell}$ is replaced with a mixture of Beta random variables of weights proportional to $p_s - p_{s+1}$, which explains the monotonicity condition on the p_s. The rest of the proof follows similarly to the previous section. Because the quantile function of mixtures cannot be directly expressed nicely, the final integral equation is messier. Nonetheless, the mixture inherits the monotonicity in n from $\beta_{n,n-s}$, and thus the uniform convergence follows. See the full proof in Section 5.9.7. □

As a special case, when the distribution $\mathbf{p}$ is only supported by 1 and 2, we have $\mathrm{CR}_1 \leqslant \mathrm{CR}(\mathbf{p}) \leqslant \mathrm{CR}_2$, and $\mathrm{CR}(\mathbf{p})$ interpolates between those two values as a function of $\mathbf{p}$ as can be seen in Section 5.8.1. This also corresponds in this special case to the comparison against a fractional order statistic.

5.8.2 Worst-case for the prophet

The competitive ratio $\mathrm{CR}_{k,\ell}$ can be interpreted as evaluating the worst-case relative loss of the online agent having less information compared to the prophet. In most settings, the prophet performs always strictly better than the online decision maker. However, here we have a prophet which, while she has more information, can be more limited in taking decisions. Hence, the worst-case distribution *for the prophet* could be of interest. Let

$$D_{k,\ell} := \sup_{n \geqslant \max(k,\ell)} \sup_{F} \frac{\sup_{\tau_1 < \cdots < \tau_k} \mathbb{E}_{X\sim F}\left[\frac{1}{k}\sum_{i\in[k]} X_{\tau_i}\right]}{\mathbb{E}_{X\sim F}\left[\frac{1}{\ell}\sum_{i\in[\ell]} X_{(i)}\right]}. \tag{5.33}$$

We can exactly compute $D_{k,\ell}$:

Proposition 5.23. *For* $(k,\ell) \in \mathrm{b}N^2$*, we have*

$$D_{k,\ell} = \max\left\{\frac{\ell}{k}, 1\right\}. \tag{5.34}$$

Proof. For the upper-bound, if $\ell \leqslant k$:

$$\sum_{i=1}^{k} X_{\tau_i} \leqslant \sum_{i=1}^{\ell} X_{(i)} + (k-\ell)\cdot\frac{1}{\ell}\sum_{i=1}^{\ell} X_{(i)} = \frac{k}{\ell}\sum_{i=1}^{\ell} X_{(i)},$$

which yields $D_{k,\ell} \leqslant 1$ by taking the expectation. If $k \leqslant \ell$, then $\sum_{i=1}^{k} X_{\tau_i} \leqslant \sum_{i=1}^{\ell}$, hence $D_{k,\ell} \leqslant \ell/k$. Overall $D_{k,\ell} \leqslant \max\{\ell/k, 1\}$. For the matching lower bound, take F the distribution such $X = 0$ with probability $1 - 1/n^2$ and $X = 1$ otherwise. The probability for at least one X_i to be equal to 1 is of order $1/n$, and the probability for strictly more than one X_i to be equal to 1 is of order $1/(2n^2)$. Therefore the chance that a second non-zero variable appears in the sequence is negligible in front of only one appearing. The optimal policy is to select a non-zero value when it appears. Thus $\lim_{n\to\infty} \mathbb{E}[\sum_{i=1}^{k} X_{\tau_i}]/\mathbb{E}[\sum_{i=1}^{\ell} X_{(i)}] = 1$, and normalizing by k and ℓ yields the worst-case instance for the prophet when $k \leqslant \ell$. For $k \geqslant \ell$ simply take a constant distribution. ☐

5.9 Detailed proofs

5.9.1 Proof of Proposition 5.7

Let $R(q) := q\mathbb{E}[X \mid X > F^{-1}(1-q)]$ be the expected reward when rejecting values below the $1-q$ quantile of F. Through a change of variable, it can be shown that $R(q) = \int_0^q F^{-1}(\theta)d\theta$.

The first step is to express the performance of ALG_n using the fact that the q_i are independently drawn from $\psi_{\ell,n-\ell}$ truncated between ε_{i-1} and ε_i. This specific step is the same as in [Cor+17]. The expected probability of not selecting any item up to i is simply $\mathbb{E}[\prod_{j\in[i]}(1-q_j)]$, and using the independence of the q_i it can be expressed as

$$\mathbb{E}[\prod_{j\in[i]}(1-q_j)] = \prod_{j\in[i]}\mathbb{E}[(1-q_j)] = \prod_{j\in[i]}\frac{a_j}{\alpha_j}.$$

Then, using that $R(q_i)$ is the expected value of selecting or not item i with threshold q_i, we have for ALG the expected performance of the algorithm that

$$\begin{aligned}
\mathrm{ALG} &= \sum_{i=1}^n \mathbb{E}[R(q_i)\prod_{j\in[i-1]}(1-q_j)]\\
&= \sum_{i=1}^n \mathbb{E}[R(q_i)]\prod_{j\in[i-1]}\frac{a_j}{\alpha_j}\\
&= \sum_{i=1}^n \int_{\varepsilon_{i-1}}^{\varepsilon_i} R(q_i)\frac{\psi_{\ell,n-\ell}(q_i)}{\alpha_i}\prod_{j\in[i-1]}\frac{a_j}{\alpha_j}\\
&= \sum_{i=1}^n \rho_i\int_{\varepsilon_{i-1}}^{\varepsilon_i} R(q_i)\psi_{\ell,n-\ell}(q_i),
\end{aligned}$$

where $\rho_i := \alpha_i^{-1}\prod_{j\in[i-1]}a_j/\alpha_j$. Hence, we obtain that

$$\begin{aligned}
\mathrm{ALG}_n &= \sum_{i=1}^n \rho_i\int_{\varepsilon_{i-1}}^{\varepsilon_i} R(q)\psi_{\ell,n-\ell}(q)dq\\
&\geqslant \min_{i\in[n]}\rho_i\sum_{i=1}^n\int_{\varepsilon_{i-1}}^{\varepsilon_i} R(q)\psi_{\ell,n-\ell}(q)dq = \min_{i\in[n]}\rho_i\int_0^1 R(q)\psi_{\ell,n-\ell}(q)dq.
\end{aligned}$$

We now show that this last integral is actually equal to $\mathrm{OPT}_{\ell,n}/n$. This was done as a special case in [Cor+17], but for $\ell \geqslant 2$ this requires the use of the distributions of order statistics.

We recall the distribution of order statistics: the density of $X_{(i)}$ for $i \in [n]$ is

$$\frac{n!}{(n-i)!(i-1)!} f(x) F(x)^{n-i} (1-F(x))^{i-1}.$$

Hence because $\mathrm{OPT}_{\ell,n} = \sum_{i=1}^{\ell} \mathbb{E}[X_{(i)}]$ we can express $\mathrm{OPT}_{\ell,n}$ using those order statistics distributions. Using the change of variable $q = 1 - F(t)$ and doing integration by parts, we obtain

$$\begin{aligned}
\mathrm{OPT}_{\ell,n} &= \sum_{i\in[\ell]} \int_0^{\omega} \frac{n!}{(n-i)!(i-1)!} t f(t) F(t)^{n-i} (1-F(t))^{i-1} dt \\
&= \sum_{i\in[\ell]} \int_0^{1} \frac{n!}{(n-i)!(i-1)!} F^{-1}(1-q)(1-q)^{n-i} q^{i-1} dq \\
&= \sum_{i\in[\ell]} \int_0^{1} \frac{n!}{(n-i)!(i-1)!} (1-q)^{n-i-1} q^{i-2} ((n-i)q - (i-1)(1-q)) \int_0^{q} F^{-1}(\theta) d\theta dq \\
&= \sum_{i\in[\ell]} \int_0^{1} \frac{n!}{(n-i)!(i-1)!} (1-q)^{n-i-1} q^{i-2} ((n-i)q - (i-1)(1-q)) R(q) dq,
\end{aligned}$$

Exchanging sum and integral, we can observe that the sum is actually telescoping:

$$\begin{aligned}
&\sum_{i\in[\ell]} \frac{n!}{(n-i)!(i-1)!} (1-q)^{n-i-1} q^{i-2} ((n-i)q - (i-1)(1-q)) \\
&= \sum_{i=1}^{\ell} \frac{n!}{(n-i-1)!(i-1)!} (1-q)^{n-i-1} q^{i-1} - \sum_{i=2}^{\ell} \frac{n!}{(n-i)!(i-2)!} (1-q)^{n-i} q^{i-2} \\
&= \sum_{i=1}^{\ell} \frac{n!}{(n-i-1)!(i-1)!} (1-q)^{n-i-1} q^{i-1} - \sum_{i=1}^{\ell-1} \frac{n!}{(n-i-1)!(i-1)!} (1-q)^{n-i-1} q^{i-1} \\
&= \frac{n!}{(n-\ell-1)!(\ell-1)!} (1-q)^{n-\ell-1} q^{\ell-1} \\
&= n \frac{(n-1)!}{(n-\ell-1)!(\ell-1)!} (1-q)^{n-\ell-1} q^{\ell-1} \\
&= n \frac{q^{\ell-1}(1-q)^{n-\ell-1}}{B(\ell, n-\ell)} \\
&= n \psi_{\ell, n-\ell}.
\end{aligned}$$

Therefore, given this algorithm, we can already deduce that for a given $n \in \mathbf{b}N$,

$$\mathrm{CR}_{\ell}(n) \geqslant \frac{\min_{i\in[n]} \rho_i}{n}.$$

5.9.2 Proof of lemma 5.8

We first recall a property of $\beta_{\ell,n-\ell}$ which can be obtained by integration by parts.

$$\beta_{\ell+1,n-\ell}(z) = \beta_{\ell,n-\ell}(z) - \frac{z^\ell(1-z)^{n-\ell}}{kB(\ell,n-\ell)}$$
$$\beta_{\ell,n-\ell+1}(z) = \beta_{\ell,n-\ell}(z) + \frac{z^\ell(1-z)^{n-\ell}}{(n-\ell)B(\ell,n-\ell)}.$$

First we remark that we can express α_i as $\beta_{\ell,n-\ell}(\varepsilon_i) - \beta_{\ell,n-\ell}(\varepsilon_{i-1}) = b_i - b_{i-1}$. Similarly for a_i,

$$\begin{aligned}
a_i &= \int_{\varepsilon_{i-1}}^{\varepsilon_i} (1-q)\psi_{\ell,n-\ell}(q)dq = \frac{\int_{\varepsilon_{i-1}}^{\varepsilon_i} q^{\ell-1}(1-q)^{n+1-\ell}dq}{B(\ell,n-\ell)} \\
&= \frac{B(\ell,n+1-\ell)}{B(\ell,n-\ell)} \frac{\int_{\varepsilon_{i-1}}^{\varepsilon_i} q^{\ell-1}(1-q)^{n+1-\ell}dq}{B(\ell,n+1-\ell)} \\
&= \frac{\frac{\Gamma(\ell)\Gamma(n+1-\ell)}{\Gamma(n+1)}}{\frac{\Gamma(\ell)\Gamma(n-\ell)}{\Gamma(n)}}(\beta_{\ell,n+1-\ell}(\varepsilon_i) - \beta_{\ell,n+1-\ell}(\varepsilon_{i-1})) \\
&= \frac{n-\ell}{n}(\beta_{\ell,n+1-\ell}(\varepsilon_i) - \beta_{\ell,n+1-\ell}(\varepsilon_{i-1})).
\end{aligned}$$

The quantities ρ_i always satisfy a simple recurrence relation, namely that $\rho_{i+1} = a_i\rho_i/\alpha_{i+1}$. So imposing the equality of the ρ_i is equivalent to the relation $\alpha_{i+1} = a_i$. Now, using this relationship:

$$\begin{aligned}
&\alpha_{i+2} = a_{i+1} \\
&\Leftrightarrow \beta_{\ell,n-\ell}(\varepsilon_{i+2}) - \beta_{\ell,n-\ell}(\varepsilon_{i+1}) = \frac{n-\ell}{n}(\beta_{\ell,n+1-\ell}(\varepsilon_{i+1}) - \beta_{\ell,n+1-\ell}(\varepsilon_i)) \\
&\Leftrightarrow b_{i+2} - b_{i+1} = \frac{n-\ell}{n}\Big(\beta_{\ell,n-\ell}(\varepsilon_{i+1}) - \beta_{\ell,n-\ell}(\varepsilon_i) \\
&\quad + \frac{\varepsilon_{i+1}^\ell(1-\varepsilon_{i+1})^{n-\ell}}{(n-\ell)B(\ell,n-\ell)} - \frac{\varepsilon_i^\ell(1-\varepsilon_i)^{n-\ell}}{(n-\ell)B(\ell,n-\ell)}\Big) \\
&\Leftrightarrow \Delta[b_{i+1}] = \Delta[b_i] - \Big(\frac{\ell}{n}\beta_{\ell,n-\ell}(\varepsilon_{i+1}) - \frac{\varepsilon_{i+1}^\ell(1-\varepsilon_{i+1})^{n-\ell}}{nB(\ell,n-\ell)} \\
&\quad - \frac{\ell}{n}\beta_{\ell,n-\ell}(\varepsilon_{i+1}) + \frac{\varepsilon_{i+1}^\ell(1-\varepsilon_{i+1})^{n-\ell}}{nB(\ell,n-\ell)}\Big) \\
&\quad \Leftrightarrow \Delta^2[b_i] = -\frac{\ell}{n}\Delta\left[\beta_{\ell,n-\ell}(\varepsilon_i) - \frac{\varepsilon_i^\ell(1-\varepsilon_i)^{n-\ell}}{\ell B(\ell,n-\ell)}\right] \\
&\quad \Leftrightarrow \Delta^2[b_i] = -\frac{\ell}{n}\Delta[\beta_{\ell+1,n-\ell}(\varepsilon_i)]
\end{aligned}$$

$$\Leftrightarrow \; \Delta^2[b_i] = \frac{\Delta[-\ell\beta_{\ell+1,n-\ell}(\beta^{-1}_{\ell,n-\ell}(b_i))]}{n}.$$

By summing those equations, we can hence obtain an explicit recurrence relationship on b_i. Indeed, for $i \geqslant 2$:

$$\sum_{j=0}^{i-2} \Delta^2[b_j] = \sum_{j=0}^{i-2} \frac{\Delta[-\ell\beta_{\ell+1,n-\ell}(\beta^{-1}_{\ell,n-\ell}(b_j))]}{n}$$

$$\Leftrightarrow (b_i - b_{i-1}) - (b_1 - b_0) = \frac{-\ell(\beta_{\ell+1,n-\ell}(\beta^{-1}_{\ell,n-\ell}(b_i)) - \beta^{-1}_{\ell,n-\ell}(0))}{n}$$

$$\Leftrightarrow b_i = b_{i-1} - \frac{\ell}{n}\beta_{\ell+1,n-\ell} \circ \beta^{-1}_{\ell,n-\ell}(b_i) + b_1.$$

5.9.3 Proof of Proposition 5.9

Because $\beta_{\ell,n-\ell}(0) = 0$ and $\beta_{\ell,n-\ell}(1) = 1$, the problem of finding an increasing partition of ε_i which satisfies $\rho_{i+1} = \rho_i$ with $\varepsilon_0 = 0$ and $\varepsilon_n = 1$ is equivalent to finding a partition of the b_i which satisfies the recurrence relation in Equation (5.13); this is a specific instance of a discrete boundary value problem. We now prove that there exists a b_1 such that $b_n = 1$, $b_{i+1} \geqslant b_i$, and it solves the discrete boundary value problem. Remark that these conditions constrain b_i to belong to the interval $[0, 1]$: this is crucial to recover from the b_i a valid sequence of ε_i as the domain of $\beta^{-1}_{\ell,n-\ell}$ is $[0, 1]$.

First note that for $b_1 \geqslant \ell/n$ and because $\beta_{\ell+1,n-\ell}(x) \leqslant 1$, we have that

$$b_{i+1} - b_i = -\frac{\ell}{n}\beta_{\ell+1,n-\ell} \circ \beta^{-1}_{\ell,n-\ell}(b_i) + b_1 \geqslant -\frac{\ell}{n} + \frac{\ell}{n} \geqslant 0.$$

This implies that any $b_1 \geqslant \ell/n$ yields an increasing sequence of b_i, hence an increasing sequence of ε_i.

Second observe that b_n can be expressed as $n-1$ times the composition of the recurrence relation, which is continuous, so b_n is a continuous function of b_1. We will show that for $b_1 = \ell/n$ and $b_1 = (\ell+1)/n$, we end up with $b_n \leqslant 1$ and $b_n \geqslant 1$ respectively. Then by intermediate value theorem this proves the existence of a b_1 in $[\ell/n, (\ell+1)/n]$ such that $b_n = 1$, and $b_{i+1} \geqslant b_i$ by the first remark because $b_1 \geqslant \ell/n$.

For $b_1 \geqslant (\ell+1)/n$, bounding the difference $b_{i+1} - b_i$ as done above for the monotonicity yields $\Delta[b_i] \geqslant 1/n$, meaning that $b_n \geqslant 1$.

The analysis for $b_1 = \ell/n$ uses a simple argument, but needs to be treated carefully. Note that if a b_1 is selected such that b_i is increasing and for some $i_0 < n$ we have $b_{i_0} \geqslant 1$, we can disregard the recurrence after i_0 and fix in the algorithm $\varepsilon_i = 1$ for $i \geqslant i_0$. This simply corresponds to ignoring any item that comes after i_0, and still produces a valid algorithm with competitive ratio lower bounded by $\ell\rho_1/n = \ell/(\alpha_1 n) = \ell/(b_1 n)$. Take $b_1 = (\ell - \delta)/n$, with $\delta > 0$ small enough such that the sequence remain increasing. Then if the corresponding b_n is greater or equal to 1, we can recover a valid sequence of ε_i discarding items after some time if necessary. This leads to a lower bound on the competitive ratio of $\ell/(\ell - \delta) > 1$. However by considering a constant distribution, it must be that the competitive ratio is smaller than 1 in the worst case. Therefore it must be that $b_n < 1$. By taking the limit as δ goes to 0, we have $b_n \leqslant 1$ for $b_1 = \ell/n$.

5.9.4 Proof of Lemma 5.10

For the detailed proof of $\mathbb{E}[X_{\tau_X}] \geqslant \mathbb{E}[Y_{\tau_Y}]$ we defer to Appendix A.1 of [Liu+21a] (note that their notation of OPT corresponds to the optimal online algorithm).

We recall that the distribution function of the ℓ-th order statistic for a random variable with distribution F is simply $\beta_{n+1-\ell,\ell} \circ F$. If we show that $\beta_{n+1-\ell,\ell} \circ F \geqslant \beta_{2n+1-\ell,\ell} \circ \sqrt{F}$, then $Y_{(\ell)}$ stochastically dominates $X_{(\ell)}$, which implies the desired result of $\mathbb{E}[Y_{(\ell)}] \geqslant \mathbb{E}[X_{(\ell)}]$ (this is immediate from writing the expectation of a positive random variable as $\int_0^\infty 1 - F$).

Let us look at the function $h(x) := \beta_{n+1-\ell,\ell}(x) - \beta_{2n+1-\ell,\ell}(\sqrt{x})$ for $x \in [0,1]$. Clearly if we show that this function is non-negative over $[0,1]$, then so is $h \circ F$ over $\mathrm{b}R^+$. Let us consider the derivative of h using the density of the beta distribution:

$$\frac{dh(x)}{dx} = \frac{x^{n-\ell}(1-x)^{\ell-1}}{B(n+1-\ell,\ell)} - \frac{\sqrt{x}^{2n-\ell-1}(1-\sqrt{x})^{\ell-1}}{2B(2n+1-\ell,\ell)}.$$

For simplicity, let $\sqrt{x} = y \in [0,1]$ then we can rewrite the above derivative as

$$\begin{aligned}
\frac{dh(x)}{dx} &= \frac{y^{2n-2\ell}(1-y^2)^{\ell-1}}{B(n+1-\ell,\ell)} - \frac{y^{2n-\ell-1}(1-y)^{\ell-1}}{2B(2n+1-\ell,\ell)} \\
&= \frac{y^{2n-2\ell}(1-y)^{\ell-1}(1+y)^{\ell-1}}{B(n+1-\ell,\ell)} - \frac{y^{2n-\ell-1}(1-y)^{\ell-1}}{2B(2n+1-\ell,\ell)} \\
&= \frac{y^{2n-\ell-1}(1-y)^{\ell-1}}{B(n+1-\ell,n)} \left(\frac{(1+y)^{\ell-1}}{y^{\ell-1}} - \frac{B(n+1-\ell,\ell)}{2B(2n+1-\ell,\ell)} \right).
\end{aligned}$$

Hence over $[0,1]$, $h'(x) \geqslant 0$ is equivalent to

$$\left(1+\frac{1}{y}\right)^{\ell-1} \geqslant \frac{B(n+1-\ell,\ell)}{2B(2n+1-\ell,\ell)}.$$

The function on the left is equal to ∞ at $y = 0$, and is decreasing in y: this implies that the derivative is positive until some $y_0 = \sqrt{x_0}$, and then possibly negative. If a function is increasing then decreasing over an interval, then its minimum is at the endpoints of the interval. And we have that $h(0) = h(1) = 0$, so $h(x) \geqslant \min_{x' \in [0,1]} h(x') = 0$, which concludes the proof.

5.9.5 Proof of Lemma 5.11

We first recall a limiting relationship between incomplete Gamma and Beta functions. We have for $X_n \sim \mathrm{Beta}(\ell, n)$ that the random variable $Y_n = nX_n$ converges in law to $\mathrm{Gamma}(\ell)$ as n goes to ∞.

In particular this means that

$$\lim_{n\to\infty} \beta_{\ell,n-\ell}\left(\frac{z}{n}\right) = \lim_{n\to\infty} \mathbb{P}(X_n \leqslant \frac{z}{n}) = \lim_{n\to\infty} \mathbb{P}(nX_n \leqslant z) = \gamma_\ell(x),$$

where the last equality stems from the equivalence between limit in distribution and point-wise limit of the distribution function (we also indirectly use that ℓX_n goes to 0 for k fixed).

Let us show the point-wise convergence. The inverse of $\beta_{\ell,n-\ell}(\cdot/n)$ is simply $n\beta_{\ell,n-\ell}^{-1}$. Because $\beta_{\ell,n-\ell}$ is continuous, the point-wise limit of the inverse is the inverse of the point-wise limit. Hence $\lim_{n\to\infty} \beta_{\ell,n-\ell}^{-1} = \gamma_\ell^{-1}$. Rewriting the main function as $-\ell\beta_{\ell+1,n-\ell} \circ \mathrm{id}/n \circ n \cdot \mathrm{id} \circ \beta_{\ell,n-\ell}^{-1}$ we have by composition that this converges to $-\ell\gamma_{\ell+1} \circ \beta_\ell^{-1}$.

Now for the uniform convergence. First, note that because of the expansion formula, we have that $\beta_{\ell,n+1-\ell} \geqslant \beta_{\ell,n-\ell}$ which means that $\beta_{\ell,n+1-\ell}^{-1} \leqslant \beta_{\ell,n-\ell}^{-1}$ (interpolating over n if necessary) so $\beta_{\ell,n-\ell}^{-1}$ decreasing in n, and $\beta_{\ell+1,n+1-\ell} \geqslant \beta_{\ell,n-\ell}$ so $\beta_{\ell+1,n+1-\ell}^{-1}$ increasing in n. This implies that $\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}$ is decreasing in n, and finally $-\ell\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}$ is increasing in n. Because this function is continuous and takes values in $[0,1]$ which is compact, we can apply Dini's Theorem which guarantees uniform convergence.

Remark. Note that we can similarly express this ODE using $\nu = \gamma_\ell^{-1}(b) \in [0, \infty)$ to avoid the use of an inverse functions:

$$\frac{d^2}{dt^2}(\gamma_\ell(\nu)) = -\ell \frac{d}{dt}(\gamma_{\ell+1}(\nu)).$$

It is not immediately clear what ν represents compared to the ε_i. Indeed, we had that $\beta_{\ell,n-\ell}^{-1}(b_i) = \varepsilon_i$, but ν cannot directly translate into ε_i as $\nu \in [0, \infty)$ whereas $\varepsilon_i \in [0, 1]$. Actually, we can definite $\nu_i = n\varepsilon_i$, which implies that

$$\Delta^2[\beta_{\ell,n-\ell}(\frac{\nu_i}{n})] = \frac{\Delta[-\ell\beta_{\ell+1,n}(\frac{\nu_i}{n})]}{n},$$

which converges using the exact same limiting argument for the pointwise limit as above to the second order ODE which $\nu(t)$ obeys. So ν is the limit of $n\varepsilon_i$. This is to put in perspective to the limit in [Cor+17] where the ODE concerned used $y_i = (1 - \varepsilon_i)^{n-1}$.

5.9.6 Proof of Theorem 5.12

Let $\xi_{\ell,n}(x) = -\ell\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(x)$ and $\xi_\ell(x) = -\ell\gamma_{\ell+1} \circ \gamma_\ell^{-1}(x)$. First

$$b_{i+1} - b_i = \frac{\xi_\ell(b_i) + c_{1,\ell}}{n} + \frac{c_{1,\ell}(n) - c_{1,\ell}}{n} + \frac{\xi_{\ell,n}(b_i) - \xi_\ell(b_i)}{n}$$

Now let us divide both side by $\xi_\ell(b_i) + c_{1,\ell}$ (which is strictly positive as $c_{1,\ell}$ must be strictly greater than ℓ)

$$\frac{c_{1,\ell}(n) - c_{1,\ell}}{n(\xi_\ell(b_i) + c_{1,\ell})} = -\frac{1}{n} + \frac{b_{i+1} - b_i}{\xi_\ell(b_i) + c_{1,\ell}} - \frac{\xi_{\ell,n}(b_i) - \xi_\ell(b_i)}{n(\xi_\ell(b_i) + c_{1,\ell})}$$

Remark also that for $x \in [0, 1]$, $x \mapsto -\ell x + \zeta_\ell(x) + c_{1,\ell}$ is bounded between $c_{1,\ell} - \ell > 0$ and $c_{1,\ell}$, therefore,

$$\left|\sum_{i=0}^{n-1} \frac{c_{1,\ell}(n) - c_{1,\ell}}{n(\xi_\ell(b_i) + c_{1,\ell})}\right| \geqslant \delta|c_{1,\ell}(n) - c_{1,\ell}|, \quad \delta > 0.$$

Now using the previous equation,

$$\left|\sum_{i=0}^{n-1} \frac{c_{1,\ell}(n) - c_{1,\ell}}{n(\xi_\ell(b_i) + c_{1,\ell})}\right| \leqslant \left|-1 + \sum_{i=0}^{n-1} \frac{b_{i+1} - b_i}{\xi_\ell(b_i) + c_{1,\ell}}\right| + \frac{1}{n(c_{1,\ell} - \ell)} \sum_{i=0}^{n-1} |\xi_{\ell,n}(b_i) - \xi_\ell(b_i)|$$

$$\leqslant \left| -1 + \sum_{i=0}^{n-1} \frac{b_{i+1} - b_i}{\xi_\ell(b_i) + c_{1,\ell}} \right| + \frac{1}{(c_{1,\ell} - \ell)} \|\xi_{\ell,n} - \xi_\ell\|_\infty.$$

The last term goes to 0 by uniform convergence, and the first one by Riemann sum and the integral equation condition, as due to $c_{1,\ell}(n)$ being bounded independently of n we have $b_{i+1} - b_i = O(1/n)$.

All in all $\lim_{n\to\infty} c_{1,\ell}(n) = c_{1,\ell}$.

5.9.7 Proof of Proposition 5.22

For this proof, we will only detail the parts which need special care compared to the original setting. Most arguments follow through identically.

First we will show that

$$\frac{1}{p_1}\mathbb{E}[X_{(S)}] = n\mathbb{E}_{q\sim\varphi_{n,\mathbf{p}}}[R(q)],$$

with q a mixture of beta random variables with distribution

$$\varphi_{n,\mathbf{p}} = \sum_{s\in[\ell]} \frac{p_s - p_{s+1}}{p_1} \cdot \beta_{s,n-s},$$

where $p_{\ell+1} = 0$. This is a proper mixture as $\sum_{s\in[\ell]}(p_s - p_{s+1})/p_1 = p_1/p_1 = 1$. This result is immediate from following the final computations of Proposition 5.7 and including the p_i:

$$\begin{aligned}
&\sum_{s\in[\ell]} \frac{p_s}{p_1} \frac{n!}{(n-s)!(s-1)!}(1-q)^{n-s-1}q^{s-2}((n-s)q - (s-1)(1-q)) \\
&= \sum_{s=1}^{\ell} \frac{p_s}{p_1} \frac{n!}{(n-s-1)!(s-1)!}(1-q)^{n-s-1}q^{s-1} - \sum_{s=2}^{\ell} \frac{p_s}{p_1} \frac{n!}{(n-s)!(s-2)!}(1-q)^{n-i}q^{s-2} \\
&= \sum_{s=1}^{\ell} \frac{p_s}{p_1} \frac{n!}{(n-s-1)!(s-1)!}(1-q)^{n-s-1}q^{s-1} - \sum_{s=1}^{\ell-1} \frac{p_{s+1}}{p_1} \frac{n!}{(n-s-1)!(s-1)!}(1-q)^{n-s-1}q^{s-1} \\
&= \sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \frac{n!}{(n-s-1)!(s-1)!}(1-q)^{n-s-1}q^{s-1} \\
&= n\sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \cdot \psi_{s,n-s} = n\varphi'_{n,\mathbf{p}},
\end{aligned}$$

as we have $\beta'_{s,n-s} = \psi_{s,n-s}$. The condition $p_s \geqslant p_{s+1}$ is clear from the above computation.

We now derive the new difference equation. First, $\alpha_i = \varphi_{n,\mathbf{p}}(\varepsilon_i) - \varphi_{n,\mathbf{p}}(\varepsilon_{i-1})$, and second

$$\begin{aligned}
a_i &= \int_{\varepsilon_{i-1}}^{\varepsilon_i} (1-q)\varphi_{n,\mathbf{p}}(q)dq \\
&= \sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \int_{\varepsilon_{i-1}}^{\varepsilon_i} (1-q)\psi_{s,n-s}(q)dq \\
&= \sum_{i=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \frac{n-s}{n} (\beta_{s,n+1-s}(\varepsilon_i) - \beta_{s,n+1-s}(\varepsilon_{i-1})
\end{aligned}$$

Let $b_i = \varphi_{n,\mathbf{p}}(\varepsilon_i)$. Because $\beta_{i,n-i}$ is an increasing bijection from $[0,1]$ to $[0,1]$, so is the mixture $\varphi_{n,\mathbf{p}}$, therefore $\varphi_{n,\mathbf{p}}^{-1}$ is well defined, and the boudnary conditions on ε_i translate to the same boundary conditions on b_i. The equation $\alpha_{i+2} = a_{i+1}$ which equalizes the ρ_i yields

$$\begin{aligned}
\Delta b_{i+1} &= \sum_{i=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \frac{n-i}{n} (\beta_{s,n+1-s}(\varepsilon_{i+1}) - \beta_{s,n+1-s}(\varepsilon_i) \\
\Leftrightarrow \Delta b_{i+1} &= \sum_{i=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \frac{n-s}{n} \Big(\beta_{s,n-s}(\varepsilon_{i+1}) - \beta_{s,n-s}(\varepsilon_i) + \frac{\varepsilon_{i+1}^s (1-\varepsilon_{i+1})^{n-s}}{(n-s)B(s,n-s)} - \frac{\varepsilon_i^s (1-\varepsilon_i)^{n-s}}{(n-s)B(s,n-s)} \Big) \\
\Leftrightarrow \Delta b_{i+1} &= \Delta \varphi_{n,\mathbf{p}}(\varepsilon_i) - \sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \frac{s}{n} \Delta \left[\beta_{s+1,n-s}(\varepsilon_i) \right] \\
\Leftrightarrow \Delta^2 b_i &= -\frac{1}{n} \Delta \left[\sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \cdot s \cdot \beta_{s+1,n-s}(\varphi_{n,\mathbf{p}}^{-1}(b_i)) \right].
\end{aligned}$$

This function, and the inverse of $\varphi_{n,\mathbf{p}}$ as a mixture inherit the monotonicity properties of $\beta_{s,n-s}$ in n, thus the point-wise convergence towards a mixture of $\mathrm{Gamma}(s)$ and its uniform convergence can be proven all the same. Let $\varphi_{\mathbf{p}}$ be the rescaled limit mixture:

$$\varphi_{\mathbf{p}} = \sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \gamma_s.$$

The difference equation yields in the limit as $n \to \infty$ the differential equation

$$b'(t) = c - \sum_{s=1}^{\ell} \frac{p_s - p_{s+1}}{p_1} \cdot s \cdot \gamma_{s+1} \circ \varphi_{\mathbf{p}}^{-1}(b(t)).$$

The rest of the proof is identical and uses the same arguments as for the case $\mathbf{p} = \mathbf{1}_\ell/\ell$. Therefore the $c(\mathbf{p})$ which satisfies the integral equation will yield the lower bound on the competitive ratio of $1/(p_\ell \cdot c(\mathbf{p}))$.

5.9.8 Proof of Proposition 5.13

The skeleton of the proof is the same as in Section 3.5 of [Liu+21a], so we defer to it for some technical details and focus here on the main differences.

We know from [Liu+21a] (Equation 23) that the optimal policy verifies

$$r'(t) = \int_{r(t)}^{\infty} \log(F(u))du, \quad r(1) = 0. \tag{5.35}$$

We will verify that the $r(t)$ proposed above is indeed the optimal policy. For $r(t) \in [q, 1]$:

$$\begin{aligned}
\int_{r(t)}^{\infty} \log(F(u))du &= \int_{r(t)}^{r(q)} \log(F(u))du + \int_{r(q)}^{\infty} \log(F(u))du \\
&= -\int_q^t \log(F(r(u))r'(u)du + (H - r(q)) \log(p) \\
&= -\int_q^t \frac{-\gamma_\ell^{-1}(1 - y(u))}{y'(u)}du + \frac{1}{y'(q)} \\
&= -\int_q^t \frac{y''(u)}{(y'(u))^2}du + \frac{1}{y'(q)} \\
&= \frac{1}{y'(t)} = r'(t),
\end{aligned}$$

where we used that $y'' = -\gamma_\ell^{-1}(1 - y) \cdot y'$. And it is optimal over $[0, q]$ as $r(t) = H$, and F_q has an atom at H, which is also the highest value.

Let us now compare $\mathbb{E}[X_\tau]$ to $\mathrm{OPT}_{\ell,n}$. The probability of a variable being greater than H and arriving in $[0, q]$ is $q \cdot (1 - F_q^{1/n}(H))$, so the probability that there is at least one variable greater than H arriving in $[0, q]$ is

$$1-(1-q(1-F_q^{1/n}(H)))^n = 1-\left(1 + q\frac{\log(F_q(H))}{n} + o(n^{-1})\right)^n \xrightarrow[n\to\infty]{} 1-\exp(q\log(F_q(H))) = 1-p^q.$$

Therefore

$$\mathrm{ALG} = (1 - p^q)H - p^q \int_q^1 \frac{1}{y'(t)}dt \xrightarrow[q\to 0]{} -\int_0^1 \frac{1}{y'(t)}dt,$$

where the last inequality follows as when $q \to 0$ then $p \to 1$, $H\log(p) \to 1/y'(0)$, and $(1-p^q)H \approx -\log(1-(1-p^q))H = -q\log(p)H \to 0$.

Let us now compute $\lim_{n\to\infty} \mathrm{OPT}_{\ell,n}$.

$$
\begin{aligned}
&\mathrm{OPT}_{\ell,n} = \sum_{i=1}^{\ell}\int_0^{\infty} 1-\sum_{j=0}^{i}\binom{n}{j}(F^{1/n}(u))^{n-j}(1-F^{1/n}(u))^j du \\
&\xrightarrow[n\to\infty]{} \sum_{i=1}^{\ell}\int_0^{\infty} 1-\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}F(u)\cdot\log^j(F(u))du \\
&= \int_0^{r(q)} \ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}F(u)\cdot\log^j(F(u))du + \int_{r(q)}^{\infty} \ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}F(u)\cdot\log^j(F(u))du \\
&= -\int_q^1 \left(\ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}F(r(t))\cdot\log^j(F(r(t)))\right)\cdot r'(t)dt \\
&+\left(\ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}p\cdot\log^j(p)\right)(H-r(q)) \\
&= -\int_q^1 \frac{\ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}F(r(t))\cdot\log^j(F(r(t)))}{y'(t)}dt \\
&+\left(\ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}p\cdot\log^j(p)\right)(H-r(q)) \\
&= -\int_q^1 \frac{\ell-\exp(-\gamma_\ell^{-1}(1-y(t)))\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^j}{j!}}{y'(t)}dt \\
&+\left(\ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}p\cdot\log^j(p)\right)(H-r(q))
\end{aligned}
$$

We now compute the limit as $q \to 0$ of the right hand term.

$$
\begin{aligned}
\left(\ell-\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{(-1)^j}{j!}p\cdot\log^j(p)\right)(H-r(q)) &= (\ell-\ell p+(\ell-1)p\log(p)+o(\log(p))\,(H-r(q)) \\
&= (-\ell\log(p)+(\ell-1)p\log(p)+o(\log(p)))\,(H-r(q)) \\
&\xrightarrow[q\to 0]{} -\frac{1}{y'(0)}.
\end{aligned}
$$

Before computing the limit, let us re-arrange the integrand of the left-hand term. Recall that $\gamma_\ell(x) = 1 - \exp(-x)\sum_{j=0}^{\ell-1} x^j/j!$.

$$
\begin{aligned}
&1 - \frac{\exp(-\gamma_\ell^{-1}(1-y(t)))}{\ell}\sum_{i=1}^{\ell}\sum_{j=0}^{i-1}\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^j}{j!} \\
&= 1 - \frac{\exp(-\gamma_\ell^{-1}(1-y(t)))}{\ell}\sum_{j=0}^{\ell-1}\sum_{i=j+1}^{\ell}\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^j}{j!} \\
&= 1 - \frac{\exp(-\gamma_\ell^{-1}(1-y(t)))}{\ell}\sum_{j=0}^{\ell-1}(\ell-j)\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^j}{j!} \\
&= \gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t)) + \frac{\exp(-\gamma_\ell^{-1}(1-y(t)))}{\ell}\left(\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^\ell}{(\ell-1)!} + \sum_{j=0}^{\ell-1} j\cdot\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^j}{j!}\right) \\
&= \gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t)) + \frac{\exp(-\gamma_\ell^{-1}(1-y(t)))}{\ell}\sum_{j=1}^{\ell}\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^j}{(j-1)!} \\
&= \gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t)) + \frac{\exp(-\gamma_\ell^{-1}(1-y(t)))}{\ell}\gamma_\ell^{-1}(1-y(t))\sum_{j=0}^{\ell-1}\frac{\left(\gamma_\ell^{-1}(1-y(t))\right)^j}{j!} \\
&= \gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t)) + \frac{\gamma_\ell^{-1}(1-y(t))}{\ell}\left(1-\gamma_\ell\circ\gamma_\ell^{-1}(1-y(t))\right) \\
&\gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t)) + \frac{\gamma_\ell^{-1}(1-y(t))}{\ell}y(t).
\end{aligned}
$$

Therefore in the limit,

$$
\lim_{q\to 0}\lim_{n\to\infty}\frac{X_\tau}{\mathrm{OPT}_{\ell,n}} = \frac{y'(0)\int_0^1 \frac{dt}{y'(t)}}{1+\ell\cdot y'(0)\int_q^1 \frac{\gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t))+\frac{\gamma_\ell^{-1}(1-y(t))}{\ell}y(t)}{y'(t)}dt}.
$$

We want to show that this quantity is equal to $1/c_\ell$. This is equivalent, after re-arranging the terms and using that $y'(0) = -c_\ell$ to

$$
\int_0^1 \frac{1}{y'(t)}\left(-c_\ell - \frac{y'}{c_\ell} + \ell\cdot\gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t)) + \gamma_\ell^{-1}(1-y(t))\cdot y(t)\right)dt = 0.
$$

Using that $y'(t) = \gamma_{\ell+1}\circ\gamma_\ell^{-1}(1-y(t)) - c_\ell$, this is equivalent to

$$
1 - \frac{1}{c_\ell} = \int_0^1 \frac{-\gamma_\ell^{-1}(1-y(t))\cdot y(t)}{y'(t)}dt.
$$

We now apply the change of variable $y = y(t)$ and $u = \gamma_\ell^{-1}(1-y)$ to the integral:

$$\int_0^1 \frac{-\gamma_\ell^{-1}(1-y(t)) \cdot y(t)}{y'(t)} dt = \int_0^\infty \frac{u \cdot (1-\gamma_\ell(u)) \cdot \gamma_\ell'(u)}{(c_\ell - \ell\gamma_{\ell+1}(u))^2} du.$$

Finally, remark that $u \cdot \gamma_\ell'(u)/(c_\ell - \ell\gamma_{\ell+1}(u))^2$ is the derivative of $1/(c_\ell - \ell\gamma_{\ell+1}(u))$. Integrating by parts, we get

$$\int_0^\infty \frac{u \cdot (1-\gamma_\ell(u)) \cdot \gamma_\ell'(u)}{(c_\ell - \ell\gamma_{\ell+1}(u))^2} du = \left[\frac{1-\gamma_\ell(u)}{c_\ell - \ell\gamma_{\ell+1}(u)}\right]_0^\infty + \int_0^\infty \frac{\gamma_\ell'(u)}{c_\ell - \ell\gamma_{\ell+1}(u)} du = -\frac{1}{c_\ell} + 1,$$

where we used in the last equality the integral characterization of c_ℓ. This concludes the proof.

5.9.9 Proof of Proposition 5.14

First let us consider the performance of the algorithm, conditionally on the q_i^j being already drawn. Basically, if while waiting to select the j-th item the algorithm arrives at step i, then the expected reward received taking into account the probability of actually selecting the item is $R(q_i^j)$. However, the probability of arriving at step i while waiting for item j is more complicated to express. Indeed, first the $j-1$ must be selected before time i, and then no item must be selected until i while waiting for j. Hence the expected reward at step i, when selecting the j-the item, can be expressed as

$$\mathbb{E}[R(q_i^j)] \sum_{t_{j-1}=j-1}^{i-1} \sum_{t_{j-2}=j-2}^{t_{j-1}-1} \cdots \sum_{t_1=1}^{t_2-1} \left[\prod_{s=1}^{j-1} \left(1 - \frac{a_{t_s}^s}{\alpha_{t_s}^s}\right) \prod_{r_1=1}^{t_1-1} \frac{a_{r_1}^1}{\alpha_{r_1}^1} \cdots \prod_{r_j=t_{j-1}+1}^{i-1} \frac{a_{r_j}^j}{\alpha_{r_j}^j}\right],$$

where the t_m correspond to the time when item m is selected, and the sums consider all possible times of selection. While this equation seems complicated, it is merely due to the fact that the thresholds depend on i and j, and all possible sequences of selection must be considered.

The proposition can then be easily obtained by re-using the fact proved in Proposition 5.7 that $\mathrm{OPT}_{\ell,n} = \mathbb{E}_{q\sim\psi_{\ell,n-\ell}}[R(q)]$ and that for any $j \in [k]$, the $(\varepsilon_i^j)_{i\in\{j-1,\ldots,n\}}$ are constructed to form a partition of $[0,1]$.

5.9.10 Proof of Lemma 5.15

Let us write down a recursion formula in i on ρ_i^j that might depend on previous values of j. We have already dealt with the case $j = 1$ when only one item could be selected by the decision maker. As such, we will assume that $j > 1$. Due to $\alpha_{i+1}^j \rho_{i+1}^j$ being equal to the probability of reaching time $i+1$ with exactly $j-1$ items already selected, denoting t_s the time when item s is selected, we have

$$\begin{aligned}
\alpha_{i+1}^j \rho_{i+1}^j &= \sum_{t_{j-1}=j-1}^{i} \sum_{t_{j-2}=j-2}^{t_{j-1}-1} \cdots \sum_{t_1=1}^{t_2-1} \mathbb{P}(\text{Selecting item } s \text{ at time } t_s \text{ for } s \in [j-1]) \\
&= \sum_{t_{j-1}=j-1}^{i} \sum_{t_{j-2}=j-2}^{t_{j-1}-1} \cdots \sum_{t_1=1}^{t_2-1} \left[\prod_{s=1}^{j-1} \left(1 - \frac{a_{t_s}^s}{\alpha_{t_s}^s}\right) \prod_{r_1=1}^{t_1-1} \frac{a_{r_1}^1}{\alpha_{r_1}^1} \cdots \prod_{r_j=t_{j-1}+1}^{i} \frac{a_{r_j}^j}{\alpha_{r_j}^j} \right] \\
&= \sum_{t_{j-1}=j-1}^{i-1} \sum_{t_{j-2}=j-2}^{t_{j-1}-1} \cdots \sum_{t_1=1}^{t_2-1} \left[\prod_{s=1}^{j-1} \left(1 - \frac{a_{t_s}^s}{\alpha_{t_s}^s}\right) \prod_{r_1=1}^{t_1-1} \frac{a_{r_1}^1}{\alpha_{r_1}^1} \cdots \prod_{r_j=t_{j-1}+1}^{i} \frac{a_{r_j}^j}{\alpha_{r_j}^j} \right] \\
&+ \sum_{t_{j-2}=j-2}^{i-1} \cdots \sum_{t_1=1}^{t_2-1} \left[\prod_{s=1}^{j-1} \left(1 - \frac{a_{t_s}^s}{\alpha_{t_s}^s}\right) \prod_{r_1=1}^{t_1-1} \frac{a_{r_1}^1}{\alpha_{r_1}^1} \cdots \prod_{r_j=i+1}^{i} \frac{a_{r_j}^j}{\alpha_{r_j}^j} \right] \\
&= \frac{a_i^j}{\alpha_i^j} \sum_{t_{j-1}=j-1}^{i-1} \sum_{t_{j-2}=j-2}^{t_{j-1}-1} \cdots \sum_{t_1=1}^{t_2-1} \left[\prod_{s=1}^{j-1} \left(1 - \frac{a_{t_s}^s}{\alpha_{t_s}^s}\right) \prod_{r_1=1}^{t_1-1} \frac{a_{r_1}^1}{\alpha_{r_1}^1} \cdots \prod_{r_j=t_{j-1}+1}^{i-1} \frac{a_{r_j}^j}{\alpha_{r_j}^j} \right] \\
&+ \left(1 - \frac{a_i^{j-1}}{\alpha_i^{j-1}}\right) \sum_{t_{j-2}=j-2}^{i-1} \cdots \sum_{t_1=1}^{t_2-1} \left[\prod_{s=1}^{j-2} \left(1 - \frac{a_{t_s}^s}{\alpha_{t_s}^s}\right) \prod_{r_1=1}^{t_1-1} \frac{a_{r_1}^1}{\alpha_{r_1}^1} \cdots \prod_{r_{j-1}=t_{j-2}+1}^{i-1} \frac{a_{r_{j-1}}^j}{\alpha_{r_{j-1}}^j} \right] \\
&= \alpha_i^j \rho_i^j + (\alpha_i^{j-1} - a_i^{j-1}) \rho_i^{j-1}.
\end{aligned}$$

Notably, if we let the ε_i^j be such that $\rho_i^j = \rho_j^j$ for all $i \in \{j, \dots, n\}$ (starting from $j = 1$, and sequentially imposing the boundary values), then dividing both sides by ρ_j^j we obtain

$$\alpha_{i+1}^j = a_i^j + (\alpha_i^{j-1} - a_i^{j-1}) \frac{\rho_{j-1}^{j-1}}{\rho_j^j}.$$

this overall yields a grid of recurrence equations for $j > 1$ and $i \geqslant j$. Now let us translate this recurrence relationship into one over b_i^j. Let us assume that $\rho_{i+1}^{j-1} = \rho_i^j$ and see what this implies for j, with $j = 1$ being already treated in Equation (5.13).

Following the same computations done in Lemma 5.8, we have that $\alpha_{i+2}^j - a_{i+1}^j = \text{cst}$ is equivalent to $\Delta^2[b_i^j] = \frac{\Delta[-\ell\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(b_i^j)]}{n} + \text{cst}$. Then, using that $\alpha_{i+1}^{j-1} - a_{i+1}^{j-1} = \alpha_{i+2}^{j-1} - \alpha_{i+2}^{j-1} + \alpha_{i+1}^{j-1} - a_{i+1}^{j-1} = -(\alpha_{i+2}^{j-1} - \alpha_{i+1}^{j-1})$, we obtain

$$\Delta^2[b_i^j] = \frac{\Delta[-\ell\beta_{\ell+1,n-\ell}\circ\beta_{\ell,n-\ell}^{-1}(b_i^j)]}{n} - \frac{\rho_{j-1}^{j-1}}{\rho_j^j}\Delta^2[b_i^{j-1}]$$
$$= \frac{\Delta[\frac{\rho_{j-1}^{j-1}}{\rho_j^j}\ell\beta_{\ell+1,n-\ell}\circ\beta_{\ell,n-\ell}^{-1}(b_i^{j-1}) - \ell\beta_{\ell+1,n-\ell}\circ\beta_{\ell,n-\ell}^{-1}(b_i^j)]}{n}.$$

Now summing these equations yields the desired result.

5.9.11 Proof of Proposition 5.16

We outline here the main steps for the proof: First, it must be that $b_j^j = o(1)$, as otherwise, because we add b_j^j for $n-j$ steps, $b_n^j = \Omega(n) > 1$. Then we show that if $\theta_{j,\ell}(n) := \rho_{j-1}^{j-1}/\rho_j^j$ is larger than some positive constant (independent of n) then $b_n^j > 1$, and if it is smaller than some positive constant then $b_n^j < 1$. By the intermediate value theorem, as b_n^j is a continuous function of b_j^j, there must be a value b_j^j such that $b_n^j = 1$, which proves the existence. Moreover, by giving an asymptotic expression of $\theta_{j,\ell}(n)$ in terms of b_j^j and b_{j-1}^{j-1}, and because $\theta_{j,\ell}(n)$ must remain bounded between two positive constants, we can inductively give an asymptotic expression of b_j^j in terms of n as $b_j^j = \Theta(n^{r_{j,\ell}})$ with $r_{j,\ell} = \ell\cdot(1-(\frac{\ell+1}{\ell})^j)$. Then we show that the b_i^j remain between 0 and 1, allowing us to map back the b_i^j solution to a solution on the ε_i^j. Finally, by evaluating the sign of ρ_i^j if any ε_{i+1}^j were to be smaller than an ε_i^j, we can obtain a contradiction, thus implying that the ε_i^j must be increasing in i, and so does the b_i^j. All of the above will be proven inductively for each j, so the aforementioned properties will be assumed true for $j-1$, and the initialization to $j=1$ already corresponds to Proposition 5.9.

Upper bound on $\theta_{j,\ell}(n)$. Let us inductively show that $\theta_{j,\ell}(n)$ is bounded where we temporarily re-define (only in this proof!) $\theta_{1,\ell}(n) := (\ell+1)/\ell$. Using the recurrence relation from Lemma 5.15, we have that

$$\begin{aligned}
\Delta[b_i^j] &\geqslant \frac{1}{n}\left(\theta_{j,\ell}(n)\ell\beta_{\ell+1}\circ\beta_\ell^{-1}(b_i^{j-1}) - \ell\right)\\
\implies b_n^j + o(1) = b_n^j - b_j^j = \frac{1}{n}\sum_{i=j}^{n-1}\Delta[b_i^j] &\geqslant \ell\left(\theta_{j,\ell}(n)\left(\frac{1}{n}\sum_{i=j}^{n-1}\beta_{\ell+1,n-\ell}\circ\beta_{\ell,n-\ell}^{-1}(b_i^{j-1})\right) - 1\right)\\
&\geqslant \ell\left(\theta_{j,\ell}(n)\left(\frac{1}{n}\sum_{i=j}^{n-1}\gamma_{\ell+1}\circ\gamma_\ell^{-1}(b_i^{j-1}) + o(1)\right) - 1\right),
\end{aligned}$$

where the last equality is due to the uniform convergence in Lemma 5.11. We now need to lower bound the above sum. Due to the induction hypothesis, we have that $b_n^{j-1} = 1$, $b_j^j = o(1/n)$ for $j > 1$ and thus $\Delta[b_i^{j-1}] \leqslant \theta_{j-1,\ell}(n)\ell/n + o(1/n)$ (this is why we define $\theta_{1,\ell}(n) = (\ell+1)/\ell$ to make sure that this expression remains true for $j = 1$ as $\Delta[b_i^1] \leqslant (\ell+1)/n$). Therefore it takes some time to go from any value $y \in (0,1)$ to $1 = b_{n-1}^j$. More specifically, if we denote t the first time where $b_t^{j-1} \geqslant y$ we obtain an inequality on t:

$$\begin{aligned}
1 = b_t^{j-1} + \sum_{i=t}^{n-1} \Delta[b_i^{j-1}] \leqslant y + \frac{1}{n}\theta_{j-1,\ell}(n)\ell + \frac{n-t-1}{n}\theta_{j-1,\ell}(n)\ell + o(1) \\
\implies n - t \geqslant \frac{1-y}{\theta_{j-1,\ell}(n)\ell} + o(1).
\end{aligned}$$

Due to the monotonicity of b_i^{j-1} in i, for any $i \geqslant t$, $b_i^{j-1} \geqslant y$. We can now lower bound the sum:

$$b_n^j \geqslant \ell\left(\theta_{j,\ell}(n)\frac{(1-y)\gamma_{\ell+1}\circ\gamma_\ell^{-1}(y)}{\ell\theta_{j-1,\ell}(n)} - 1\right) + o(\theta_{j,\ell}(n)). \tag{5.36}$$

We could get tighter bounds on $\theta_{j,\ell}(n)$ if we maximize this inequality in y, but let us simply take $y = 1/2$, which numerically is not so bad for low values of ℓ and looks to be the maximum for large values of ℓ anyway. All in all, for b_n^j to remain below 1, it must be that

$$\theta_{j,\ell}(n) + o(\theta_{j,\ell}(n)) \leqslant \frac{2(\ell+1)}{\gamma_{\ell+1}\circ\gamma_\ell^{-1}(1/2)}\theta_{j-1,\ell}(n).$$

Iterating this inequality in j with $\theta_{1,\ell}(n)$ bounded shows that for n large enough all the $\theta_{j,\ell}(n)$ indeed remain bounded.

Size estimate of b_j^j. Before showing that $\theta_{j,\ell}(n)$ is bounded below by some positive constant, we will first show that b_j^j must be very small in front of $1/n$ due to the previous upper bound on $\theta_{j,\ell}(n)$. Let us express ρ_j^j in terms of ρ_{j-1}^{j-1}:

$$\begin{aligned}
\rho_j^j &= \frac{1}{\alpha_j^j}\prod_{r\in[j-1]}\left(1 - \frac{a_r^r}{\alpha_r^r}\right) = \frac{1}{\alpha_j^j}\left(1 - \frac{a_{j-1}^{j-1}}{\alpha_{j-1}^{j-1}}\right)\alpha_{j-1}^{j-1}\rho_{j-1}^{j-1} \\
&= \frac{1}{\alpha_j^j}\left(\alpha_{j-1}^{j-1} - a_{j-1}^{j-1}\right)\rho_{j-1}^{j-1} \\
&= \frac{\ell\beta_{\ell+1,n-\ell}\circ\beta_{\ell,n-\ell}^{-1}(b_{j-1}^{j-1})}{nb_j^j}\rho_{j-1}^{j-1},
\end{aligned}$$

where the last equality can be obtained following the same computations done in Section 5.9.2. This implies that $b_j^j = O((\ell\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}(b_{j-1}^{j-1}))/n)$.

The expansion of $\beta_{\ell,n-\ell}^{-1}(x)$ around 0 is $(\ell \cdot xB(\ell, n-\ell))^{1/\ell} + o(x^{1/\ell}) = (\ell \cdot x)^{1/\ell}/n + o(1)$. Using the combinatorial formula for $\beta_{\ell+1,n-\ell}$, we have

$$\begin{aligned}
\ell\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}(b_j^j) &= \ell \sum_{t=\ell+1}^{n} \binom{n}{t} \beta_{\ell,n-\ell}^{-1}(b_j^j)^t \cdot (1 - \beta_{\ell,n-\ell}^{-1}(b_j^j))^{n-t} \\
&= \ell \frac{n^{\ell+1}}{(\ell+1)!} \frac{(\ell! \cdot b_j^j)^{1+1/\ell}}{n^{\ell+1}} + o((b_j^j)^{1+1/\ell}) \\
&= \frac{\ell}{\ell+1} (\ell!)^{1/\ell} (b_j^j)^{1+1/\ell} + o((b_j^j)^{1+1/\ell}). \qquad (5.37)
\end{aligned}$$

The second equality is because $(1 - \beta_{\ell,n-\ell}^{-1})^{n-t} \to_{n\to\infty} 1$ as long as $b_j^j = o(1)$, and because

$$\binom{n}{t+1} \beta_{\ell,n-\ell}^{-1}(b_j^j)^{t+1} / (\binom{n}{t} \beta_{\ell,n-\ell}^{-1}(b_j^j)^t) = O((b_j^j)^\ell) = o(1),$$

so the first term of the sum dominates the other ones. Hence, using the growth rate induction hypothesis $b_{j-1}^{j-1} = \Theta(n^{r_{j-1,\ell}})$, we can upper bound b_{j-1}^{j-1} by $O(n^{(1+1/\ell)r_{j-1,\ell}-1})$. We now solve the recurrence $r_{j,\ell} = ((\ell+1)/\ell) r_{j-1,\ell} - 1$ which will allow us to conclude that $b_j^j \leqslant O(n^{r_{j,\ell}})$. The first term and the fixed point of this recurrence are respectively $r_{1,\ell} = -1$ and ℓ, a classic exercise shows that $r_{j,\ell} = -((\ell+1)/\ell)^{j-1}(1+\ell) + \ell = -\ell(((\ell+1)/\ell)^j - 1)$. Moreover, for $j > 1$, $r_{j,\ell} < -1$, meaning that $b_j^j = o(n^{-1})$.

Lower bound on $\theta_{j,\ell}(n)$. We can now proceed to lower bound $\theta_{j,\ell}(n)$. For $\Delta[b_i^j]$ to be big enough and for b_n^j to reach 1, $\theta_{j,\ell}(n)$ must be large enough. Indeed

$$\Delta[b_i^j] \leqslant \frac{\theta_{j,\ell}(n)\ell}{n} + b_j^j = \frac{\theta_{j,\ell}(n)\ell}{n} + o(\frac{1}{n}),$$

using the previous upper bound on b_j^j and that $0 \leqslant \beta_{\ell+1,n-\ell} \leqslant 1$. Therefore $b_n^j \leqslant \theta_{j,\ell}(n)\ell + o(1)$, which implies that if $\theta_{j,\ell}(n)$ is strictly smaller than $1/\ell + o(1)$, then b_n^j is smaller than 1. We can then apply the intermediate value theorem to the function $b_n^j(b_j^j)$ to obtain the existence, and for the value b_j^j which satisfies the boundary value condition, it must be that $\theta_{j,\ell}(n) \geqslant 1/\ell + o(1)$. Using once again the expression of $\theta_{j,\ell}(n)$ and induction hypothesis on b_{j-1}^{j-1}, we can conclude that $b_j^j = \Omega(n^{r_{j,\ell}})$ and therefore that $b_j^j = c_{j,\ell} n^{r_{j,\ell}}$ with $c_{j,\ell} \in [1/\ell, (2(\ell+1)/(\gamma_{\ell+1} \circ \gamma_\ell^{-1}(1/2)))^{j-1}(\ell+1)/\ell] + o(1)$. The actual constants could be tightened, and this would immediately yield a lower bound on the competitive ratio, akin to using the bound $c_{1,\ell} \leqslant \ell + 1$ to prove that $\mathrm{CR}_{1,\ell} \geqslant \ell/(\ell+1)$.

Mapping b_i^j back to ε_i^j. To ensure that the solution to the discrete boundary value problem in b_i^j translates into a solution to the discrete boundary value problem in ε_i^j, we must ensure that the b_i^j remain in $[0,1]$ for $\beta_{\ell,n-\ell}^{-1}(b_i^j)$ to be well defined. For $b_i^j \leqslant 1$, one way to see this is to define a continuous extension of $\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}(x)$ by 1 for any $x > 1$, and by 0 whenever $x < 0$. This ensures that if for some i, $b_i^j > 1$ then it remains strictly greater than 1. Indeed for t_1 the first time it crosses 1 the difference $\Delta[b_{t_1}^j]$ must be positive, and for any $t > t_1$ $\Delta[b_t^j] \geqslant \Delta[b_{t_1}^j] \geqslant 0$ due to the monotonicity of b_i^{j-1} by the induction hypothesis and due to the continuous extension which remains fixed at 1. This entails that $b_n^j > 1$. This is a contradiction with $b_n^j = 1$.

Now let us show that the b_i^j that solves the boundary value problem always remain positive. The main idea is that, due to the relative size of the b_j^j compared to the b_j^{j-1}, the sequence must be increasing at the beginning and hence positive. After some time due to the 'discrete Lipschitzness' the sequence i.e. the difference between two consecutive terms is bounded, and because the b_i^{j-1} are increasing, it cannot go below a certain positive threshold without encountering previous values of b_i^j which were increasing and thus the sequence must go back up. First, when $b_i^j = o(1)$, we can always approximate $\ell\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}(b_i^j) = L(b_i^j)^{1+1/\ell} + o((b_i^j)^{1+1/\ell})$, where $L = \frac{\ell}{\ell+1}(\ell!)^{1/\ell}$. From a high level, what this means is that whenever x is small when we compare x and $\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(x) \approx L \cdot x \cdot x^{1/\ell}$ the first will dominate the second. So, for any $i \leqslant n/\log(n)$, $\Delta[b_i^1] = c_{1,\ell}/n + o(1/n)$, and $b_i^1 = c_{1,\ell}(i/n) + o(i/n)$. From there we can verify by induction that for $i \leqslant n/\log(n)$, $b_i^j = \kappa_j \cdot (n/i)^{r_{j,\ell}} + o((n/i)^{r_{j,\ell}})$ with κ_j bounded between two positive constant independent of n, and $r_{j,\ell} = -\ell(((\ell+1)/\ell)^j - 1)$. We can start by upper bounding $\Delta[b_i^j]$ using $i \leqslant n/\log(n)$ by

$$\begin{aligned}
\Delta[b_i^j] &\leqslant \theta_{j,\ell}(n)L\frac{1}{n}(b_i^{j-1})^{1/\ell+1} + c_{j,\ell}(n)n^{r_{j,\ell}} \\
&\leqslant \theta_{j,\ell}(n)L\kappa_{j-1}^{1+1/\ell}\frac{1}{n}\left(\frac{n}{i}\right)^{(1+1/\ell)\cdot r_{j-1,\ell}} + c_{j,\ell}(n)n^{r_{j,\ell}} \\
&\leqslant \kappa_j\frac{\log(n)^{(1+1/\ell)\cdot r_{j-1,\ell}}}{n} + o(\frac{\log(n)^{(1+1/\ell)\cdot r_{j,\ell}}}{n}),
\end{aligned}$$

with $\kappa_j = \theta_{j,\ell}(n)L\kappa_{j-1}^{1+1/\ell} + c_{j,\ell}(n)$. This implies that for $i \leqslant n/\log(n)$:

$$b_i^j \leqslant \kappa_j(i\log(n)^{(1+1/\ell)\cdot r_{j-1,\ell}})/n + o((i\log(n)^{(1+1/\ell)\cdot r_{j,\ell}})/n) \leqslant \kappa_j\log(n)^{r_{j,\ell}} + o(\log(n)^{r_{j,\ell}}).$$

Therefore b_i^j is negligible in front of b_i^{j-1}, so we can redo the same computations by replacing above the inequality by an equality as $\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(b_i^j)$ is negligible in front of $\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(b_i^{j-1})$. This also implies that $\Delta[b_i^j] \geqslant 0$ as the only negative term is negligible. We obtain $b_i^j = \kappa_j \cdot (n/i)^{r_{j,\ell}} + o((n/i)^{r_{j,\ell}})$. Now for $i = \lfloor \log(n)/n \rfloor$, we have that $b_i^j = \kappa_j \log(n)^{r_{j,\ell}} + o(\log(n)^{r_{j,\ell}})$, which is strictly greater than ℓ/n, an upper bound on the minimum value of $\Delta[b_i^j]$ derived from Lemma 5.15. Because $\Delta[b_i^j] \geqslant 0$ for any $i \leqslant n/\log(n)$ and b_i^{j-1} is non decreasing, for any t such that $b_i^j \leqslant b_{n/\log(n)}$ we have that $\Delta[b_t^j] \geqslant \Delta[b_{n/\log(n)}^j] \geqslant 0$. Thus the sequence b_i^j cannot keep on decreasing when going below $b_{n/\log(n)}^j$, and $b_{n/\log(n)}^j - \ell/n > 0$ which overall yields the non-negativity of the b_i^j.

Monotonicity of the b_i^j. It remains to show the monotonicity. The quantity ρ_i^j, while it stems from a probabilistic event that had assumed that the ε_i^j were increasing, can be defined for any ε_i^j using integrals with no further requirements on the ε_i^j. For now, we have shown that there exists ε_i^j such that all the ρ_i^j are equal, $\varepsilon_{j-1}^j = 0$ and $\varepsilon_n^j = 1$. First, the sign of $\alpha_i^j \rho_i^j$ is always positive. Indeed, α_i^j and a_i^j are of the same sign, positive if $\varepsilon_i^j \geqslant \varepsilon_{i-1}^j$, and negative otherwise. In both cases, the ratio is positive and smaller than 1, which also implies the positivity of $1 - a_i^j/\alpha_i^j$. As a sum of products of positive terms, $\alpha_i^j \rho_i^j$ is always positive. Because all the ρ_i^j are equal, they have the same sign, so either all the ρ_i^j and α_i^j are positive, or they are all negative. Because $\varepsilon_{j-1}^j = 0$ and $\varepsilon_n^j = 1$, there must be some $t \geqslant j$ such that $\varepsilon_t^j \geqslant \varepsilon_{t-1}^j$, implying that α_t^j is positive. Therefore all the $\alpha_i^j = \beta_{\ell,n-\ell}(\varepsilon_i^j) - \beta_{\ell,n-\ell}(\varepsilon_{i-1}^j)$ are positive, which means that the ε_i^j are non-decreasing. Finally, because $\varepsilon_{j-1}^n = 0 \neq 1 = \varepsilon_n^j$, then at least one of the ρ_i^j is finite, which by equality of the ρ_i^j means that all of them must be finite, and so all the ε_i^j are distinct. Hence the ε_i^j are non-decreasing and distinct, so are increasing.

5.9.12 Proof of Proposition 5.17

Due to the uniform convergence in Lemma 5.11, it seems intuitive that we do have the convergence from the discrete boundary value problem to the continuous one. However, there are many technical difficulties that make proving this convergence especially challenging, in particular the fact that the limit function $\gamma_{\ell+1} \circ \gamma_\ell^{-1}$ is not Lipschitz over $[0, 1]$ having an infinite derivative at 1. The convergence of $c_{1,\ell}(n)$ is already proven in Theorem 5.12, so it remains to prove the convergence of $\theta_{j,\ell}(n)$ towards $\theta_{j,\ell}$ (the solution to the continuous boundary value problem).

Let $\xi_{\ell,n}(x) = -\ell\beta_{\ell+1,n-\ell} \circ \beta_{\ell,n-\ell}^{-1}(x)$ and $\xi_\ell(x) = -\ell\gamma_{\ell+1} \circ \gamma_\ell^{-1}(x)$. First due to the boundedness of the sequence $\theta_{j,\ell}(n)$, by the Bolzano-Weierstrass theorem there is at least one subsequence with an accumulation point θ, and we will work with such a subsequence.

Existence of solution to ODE. We now consider the ODE $y'(t) = \xi_\ell(y) - \theta\xi_\ell(b^{j-1}(t))$ with initial condition $y(0) = 0$. as ξ_ℓ and $\xi_\ell(b^{j-1}(t))$ are continuous, there exists a solution b^j over $[0, 1]$. We also have that

$$\xi_\ell'(x) = -\ell \cdot (\gamma_\ell^{-1})'(x) \cdot (\gamma_{\ell+1})'(\gamma_\ell^{-1}(x)) = -\ell \cdot \frac{\gamma_{\ell+1}' \circ \gamma_\ell^{-1}(x)}{\gamma_\ell' \circ \gamma_\ell^{-1}(x)} = -\gamma_\ell^{-1}(x),$$

so as long as $x < 1$ then ξ_ℓ' is bounded and $\gamma_\ell^{-1}(x)$ Lipschitz over the interval $[0, x]$. This means that as long as b^j is strictly smaller than 1, then by the Cauchy Lipschitz theorem the solution must be unique. The case when $\max_{[0,1]} b_j < 1$ is easier to treat, so we focus on when $\max_{[0,1]} b_j \geqslant 1$. Because ξ_ℓ is bounded, b^j itself is Lipschitz over $[0, 1]$, so denoting t_1 the first time $b^j = 1$ the solution is unique over $[0, t_1]$ due to the Lipschitzness of b^j. Note that over $(t_1, 1]$ there can be potentially multiple solutions satisfying the initial condition $b^j(0) = 0$.

We now wish to prove the convergence of $b_{\lfloor tn \rfloor}^j$ towards $b^j(t)$ for any $t \in [0, t_1]$. We will prove it first on $[0, t_1 - \varepsilon]$ for any $\varepsilon > 0$. The main idea is that b_i^j is almost an Euler discretization of the continuous solution, and the same ideas used in the convergence of the Euler method can be modified to take into account that the discretization uses $\xi_{\ell,n}$ and not ξ_ℓ.

$b_{\lfloor tn \rfloor}$ **and** $b(t)$ **are different from** 1. Let $t \in [0, t_1 - \varepsilon]$, in which case $b_j(t) < 1$ by continuity as $t < t_1$ and t_1 is the first time for which $b^j(t) = 1$. Similarly, we now show that for n large enough, $b_{\lfloor tn \rfloor} \leqslant 1 - c$ for some $c > 0$. The quantity $b_{\lfloor tn \rfloor}$ is bounded between $(0, 1)$ so has a non-empty set of accumulation points. For n large enough, the distance between $b_{\lfloor tn \rfloor}$ and the set of accumulation points will go to 0. If not we can look at the sub-sequence of points which do not converge to an accumulation point and apply Bolzano-Weierstrass again to exhibit a new accumulation point towards which at least some of the points converge, showing a contradiction. Therefore, if all the accumulation points are strictly smaller than 1, then there exists some constant c such that for n large enough $b_{\lfloor tn \rfloor} \leqslant 1 - c$. All the accumulation points must be smaller than 1 as the b_i^j are smaller than 1. Suppose that 1 belongs to the set of accumulation points. Because $b_n^j = 1$, this means that $b_n^j - b_{\lfloor tn \rfloor}^j = n^{-1} \sum_{i=\lfloor tn \rfloor}^n \Delta[b_i^j] \to 0$.

So, due to the monotonicity of b_i^j and the convergence of b_i^{j-1} by the induction hypothesis, for any t' in $[t,1]$ we have $b^j_{\lfloor t'n \rfloor} \to 1$, and $n\Delta[b^j_{\lfloor t'n \rfloor}] \to \theta f(b^{j-1}(t')) - \ell$. Using that $\xi_{\ell,n}$ converges uniformly towards ξ_ℓ, the Riemann sum approximation tells us that $n^{-1}\sum_{i=\lfloor t'n \rfloor}^{n} \Delta[b_i^j] \to \int_{t'}^{1} \theta\xi_\ell(b^{j-1}(u))du - \ell = 0$. This equation is valid for any $t' > t$, which is not possible as b^{j-1} is increasing and therefore the integral value must be different. Overall, we have proven that the $b^j_{\lfloor tn \rfloor}$ must remain far from 1 as long as $t < 1$.

Convergence by Euler's method. Instead of working with the discrete sequence b_i^j, we work with the affine by parts function $b^j_{(n)}$ which takes value b_i^j at times i/n and each of those values are interpolated through linear segments. We will prove the uniform convergence of this affine by part function to the continuous limit. One way to prove this could be to use the Arzela-Ascoli theorem, as we have now a sequence of functions with approximately the same Lipschitz constant, and are thus equicontinuous. Instead we will apply Euler's method. Let $\varphi(y,t) = \xi_\ell(y) - \theta\xi_\ell(b^{j-1}(t))$ and $\varphi_n(y,t) = \xi_{\ell,n}(y) - \theta_{j,\ell}(n)\xi_{\ell,n}(b^{j-1}_{(n)}(t))$. The function φ_n converges uniformly to φ due to Lemma 5.11, that $\theta_{j,\ell}(n)$ converges to some θ for the subsequence considered, and $b^{j-1}_{(n)}$ converges uniformly to its limit solution. Let $t_i = i/n$, and $\delta_i = b^j(t_i) - b_i^j$ be the global truncation error up to time t_i (the notation is omitted but b_i^j depends on n). We only consider time t_i with $t_i \leqslant t_1 - \varepsilon$, so that $\varphi(y,t)$ is B-Lipschitz in y for $B > 0$ and $b^j(t)'' = (b^j)'(t)\xi'_\ell(b^j(t)) - \theta \cdot (b^{j-1})'(t)\xi'_\ell(b^{j-1}(t))$ is also bounded by some $A > 0$ due to the boundedness of $(b^{j-1})'$ and $(b^j)'$. Looking at the global truncation error, denoting by $\omega_n = \|\varphi_n - \varphi\|_\infty$, we have

$$
\begin{aligned}
&|\delta_{i+1}| = |b^j(t_{i+1}) - b^j_{i+1} = b^j(t_{i+1}) - b_i^j - \frac{\varphi_n(b_i^j, t_i)}{n}| \\
&= |b^j(t_{i+1}) - b_i^j - \frac{\varphi(b_i^j, t_i)}{n} + \frac{\varphi(b_i^j, t_i) - \varphi_n(b_i^j, t_i)}{n}| \\
&= |b^j(t_i) - b_i^j + \frac{\varphi(b^j(t_i), t_i) - \varphi(b_i^j, t_i)}{n} + (b^j(t_{i+1}) - b^j(t_i) - \frac{1}{n}\varphi(b^j(t_i), t_i)) + \frac{\varphi(b_i^j, t_i) - \varphi_n(b_i^j, t_i)}{n}| \\
&\leqslant |\delta_i| + \frac{1}{n}B|\delta_i| + \frac{1}{2}A\frac{1}{n^2} + \frac{\omega_n}{n} \leqslant (1 + \frac{B}{n})\delta_i + \frac{\omega'_n}{n},
\end{aligned}
$$

where $\omega'(n) = \max(A/(2n), \omega_n)$ which goes to zero as both A/n and ω_n do. Using that $\delta_0 = 0$, we can apply this inequality iteratively leading to

$$
\delta_i \leqslant \sum_{j=1}^{i-1} \left(1 + \frac{B}{n}\right)^j \frac{\omega'_n}{n} = \frac{\omega'_n}{n} \frac{(1 + B/n)^i - 1}{1 + B/n - 1} \leqslant \omega'_n \exp((t_1 - \varepsilon)B) \xrightarrow[n \longrightarrow \infty]{} 0.
$$

Because $t_{i+1} - t_i \to 0$ and by Lipschitzness of $b^j_{(n)}$ we have the convergence towards b^j for any $t \in [0, t_1 - \varepsilon]$. Finally

$$|b^j(t_1) - b^j_{(n)}(t_1)| \leqslant |b^j(t_1) - b^j(t_1 - \varepsilon)| + |b^j(t_1 - \varepsilon) - b^j_{(n)}(t_1 - \varepsilon)| + |b^j_{(n)}(t_1 - \varepsilon) - b^j_{(n)}(t_1)|$$
$$\leqslant M\varepsilon + \delta_{\lfloor (t_1 - \varepsilon)n \rfloor}(\varepsilon) + M\varepsilon,$$

with M a common upper bound on φ_n and φ for n large enough. We can take the limit of this inequality over n for any fixed n, and then take the limit over ε. This implies that $\lim_{n\to\infty} b^j_{(n)}(t_1) = b^j(t_1) = 1$, which is impossible unless $t_1 = 1$ as we have already proven that this limit is different from 1 as long as $t < 1$. This implies that θ is a solution to the continuous boundary value problem. This also proves the existence of a solution to the continuous boundary value problem.

We now prove that the solution of the continuous boundary value problem must be unique. We will show that $b^j(1)$ is strictly increasing in the parameter $\theta \geqslant$ of the ODE. Let $\theta_1 > \theta_2$, b_1 and b_2 the respective solutions, and $d = b_2 - b_1$ their difference. First we show that $b_1 \geqslant b_2$. Let $M = \max_{[0,t]} d(t)$, t_0 the point at which the maximum is reached, and $M \geqslant d(0) = b_2(0) - b_1(0) = 0$. If $t_0 > 0$, we have at t_0 that $d'(t_0) = b_2'(t_0) - b_1'(t_0) = \xi_\ell(b_2(t_0)) - \xi_\ell(b_1(t_0)) - (\theta_2 - \theta_1)\xi_\ell(b^{j-1}(t_0) < 0$ using that $b_2(t_0) \geqslant b_1(t_0)$, ξ_ℓ is decreasing, and $\xi_\ell(b^{j-1}(t_0)) < 0$ as $t_0 > 0$. By continuity d is strictly decreasing in a neighborhood of t_0, and therefore for $\delta > 0$ small enough $d(t_0 - \delta) > d(t_0) = M$ which is impossible by definition of M. Thus $t_0 = 0$ and $b_1 \geqslant b_0$. Finally if there is some $t_0 > 0$ such that $b_1(t_0) = b_2(t_0)$, the same argument yields $d' < 0$ which with $d(t_0) = 0$ contradicts $d \geqslant 0$. Hence $b_1(t) > b_2(t)$ over $(0, 1]$ and in particular $b_1(1) > b_2(1)$.

Because there is a unique possible value for the limit, there is only one possible accumulation point for $\theta_{j,\ell}(n)$, implying that $\theta_{j,\ell}(n)$ does converge to the unique solution of the continuous boundary value problem in Equation (5.6).

To finish, b^j is non-decreasing as $\Delta[b^j_i] \geqslant 0$ and it must remain so in the limit. Moreover because b^{j-1} is strictly increasing, so is b^j. The initialization for this property is that b^1 is strictly increasing as $c_{1,\ell} > 1$. Additionally the convergence of $\theta_{j,\ell}(n)$ implies the convergence of $c_{j,\ell}(n)$, and the relation between these two quantities is immediate from taking the limit in Equation (5.37). The monotonicity of b^j immediately implies that $\theta_{j,\ell} \geqslant 1$, as $(b^j)'(1) = \theta_{j,\ell} \cdot \ell - \ell = \ell(\theta_{j,\ell} - 1) \geqslant 0$.

5.9.13 Proof of Proposition 5.19

The proof consists of two steps, using Jensen's inequality on the reward of the single threshold algorithm similarly done in [Cor+19a] to prove in a simple way the $1-1/e$ performance of $F^{-1}(1-1/n)$ in the i.i.d. single item setting, and algebraic manipulations as well as inequalities to obtain the desired lower bound.

We can start by noting that the expression of the online algorithm when only a single threshold is used is much simpler. Let $q \in [0,1]$ be the quantile corresponding to the selected threshold, e.g. $T = F^{-1}(1-q)$. The expected reward given by the j-th item at time i is simply $R(q)$ times the probability of having selected exactly $j-1$ item up to time $i-1$, which corresponds to a random variable distributed according to $\mathrm{Binomial}(i-1, q)$ to be equal to $j-1$. The total expected reward obtained through the j-th item is thus

$$\sum_{i=j}^{n} R(q) \binom{i-1}{j-1} q^{j-1}(1-q)^{i-j}.$$

Moreover, we know through the proof of Proposition 5.7 that for Q distributed according to $\mathrm{Beta}(\ell, n-\ell)$, $\mathrm{OPT}_{\ell,n} = n\mathbb{E}[R(Q)]$. The expectation of Q is $\mathbb{E}[Q] = \ell/n$, and using the concavity of R we obtain

$$\mathrm{OPT}_{\ell,n} = n\mathbb{E}[R(Q)] \leqslant nR\left(\frac{\ell}{n}\right). \tag{5.38}$$

Due to this inequality, we set the deterministic quantile to be $q = \ell/n$, which immediately implies that

$$\sum_{i=j}^{n} R(q) \binom{i-1}{j-1} q^{j-1}(1-q)^{i-j} \geqslant \left(\sum_{i=j}^{n} \frac{1}{n} \binom{i-1}{j-1} \left(\frac{\ell}{n}\right)^{j-1} \left(1-\frac{\ell}{n}\right)^{i-j} \right) \mathrm{OPT}_{\ell,n}\,.$$

To obtain a lower bound on this competitive ratio, it remains to lower bound this sum, which we will denote by $S_{j,\ell,n}$.

$$S_{j,\ell,n} = \sum_{i=0}^{n-j} \frac{1}{n} \binom{i+j-1}{j-1} \left(\frac{\ell}{n}\right)^{j-1} \left(1-\frac{\ell}{n}\right)^{i} = \frac{\ell^{j-1}}{n^j (j-1)!} \sum_{i=0}^{n-j} \frac{(i+j-1)!}{i!} \left(1-\frac{\ell}{n}\right)^{i}.$$

We recognize that $\sum_{i=0}^{n-j} (i+j-1) \times \cdots \times (i+1)(1-\ell/n)^i$ is the $j-1$-th derivative of the geometric sum $\sum_{i=0}^{n-1}(1-\ell/n)^i = (1-(1-\ell/n)^n)/(1-(1-\ell/n))$. the t-th derivative of $1/(1-x)$ is $t!/(1-x)^{t+1}$ and the t-th derivative of $1-x^n$ for $t \geqslant 1$ is $-n!/(n-t)!x^{n-t}\mathbf{b}1[t \leqslant n]$. Using Leibniz rule for derivation,

$$\begin{aligned}
\frac{d^{j-1}}{dx}\frac{1-x^n}{1-x} &= \sum_{t=0}^{j-1}\binom{t-1}{t}\frac{d^{j-1-t}}{dx}(1-x^n)\frac{d^t}{dx}\left(\frac{1}{1-x}\right) \\
&= (j-1)!\frac{1-x^n}{(1-x)^j} - \sum_{t=0}^{j-2}\binom{j-1}{t}\cdot\frac{t!}{(1-x)^{t+1}}\cdot\frac{n!}{(n+1+t-j)!}x^{n+1+t-j} \\
&= (j-1)!\frac{1}{(1-x)^j} - \sum_{t=0}^{j-1}\binom{j-1}{t}\cdot\frac{t!}{(1-x)^{t+1}}\cdot\frac{n!}{(n+1+t-j)!}x^{n+1+t-j}
\end{aligned}$$

For $x = (1-\ell/n) \leqslant 1$, $x^{n+1+t-j} = (1-\ell/n)^{n+1+t-j} \leqslant \exp(-\ell)(1-\ell/n)^{1+t-j}$, $(1-x)^{t+1} = (\ell/n)^{t+1}$, and $n!/(n+1+t-j)! \leqslant n^{j-t-1}$. Therefore

$$\begin{aligned}
S_{j,\ell,n} &\geqslant \frac{\ell^{j-1}}{n^j(j-1)!}\left[(j-1)!\frac{1}{(\ell/n)^j} - \sum_{t=0}^{j-1}\frac{(j-1)!}{t!(j-1-t)!}\cdot\frac{t!}{(\ell/n)^{t+1}}\cdot n^{j-t-1}\exp(-\ell)(1-\frac{\ell}{n})^{1+t-j}\right] \\
&= \frac{1}{\ell}\left(1-e^{-\ell}\sum_{t=0}^{j-1}\frac{(\ell/(1-\ell/n))^{j-1-t}}{(j-1-t)!}\right) \\
&= \frac{1}{\ell}\left(1-e^{-\ell}\sum_{t=0}^{j-1}\frac{(\ell/(1-\ell/n))^{t}}{t!}\right) \\
&= \frac{1}{\ell}\left(1-e^{\ell^2/(n-\ell)}+e^{\ell^2/(n-\ell)}-e^{-\ell}\cdot e^{\ell/(1-\ell/n)}\cdot e^{-\ell/(1-n/\ell)}\sum_{t=0}^{j-1}\frac{(\ell/(1-\ell/n))^{t}}{t!}\right) \\
&= \frac{\gamma_j(\ell/(1-\ell/n))e^{\ell^2/(n-\ell)}+(1-e^{\ell^2/(n-\ell)})}{\ell} \geqslant \frac{\gamma_j(\ell)}{\ell} - \frac{1}{n}\left(\ell-\gamma_j(\ell)\cdot\ell-\frac{\ell^j e^{-\ell}}{(j-1)!}\right) - o\left(\frac{1}{n^2}\right),
\end{aligned}$$

where we used Taylor approximations to get estimates of $\gamma_j(\ell/(1-\ell/n))$, and $e^{\ell^2/(n-\ell)}$. Summing the contribution of every item $j \in [k]$, we immediately obtain the desired lower bound

$$\mathrm{CR}^S_{k,\ell}(n) \geqslant \frac{\sum_{j=1}^k \gamma_\ell(j)}{k} - \frac{1}{n}\left(1-\gamma_j(\ell)-\frac{\ell^{j-1}e^{-\ell}}{(j-1)!}\right) - o\left(\frac{1}{n^2}\right).$$

5.9.14 Proof of Proposition 5.20

Once the worst case instance from [AM21] is correctly modified, their proof almost entirely follows through. First of all, they show that for any quantity W independent of n, the prophet's expected reward is at least

$$\ell + W - \frac{1+W}{n+1},$$

and the decision maker's expected reward is at most

$$\mathbb{E}[\min(\text{Poisson}(nq), k)] \left(1 + \frac{W}{nq}\right) + 2kWn^{-2/3} + 2kn^{-1/3},$$

with q the probability of accepting any item which is a function of the random tie-break probability. They further show that the derivative of $\mathbb{E}[\min(\text{Poisson}(nq), k)] \, (1 + W/nq)$ in $\lambda = nq$ is equal to

$$\frac{d}{d\lambda}\mathbb{E}[\min(\text{Poisson}(\lambda), k)] \left(1 + \frac{W}{\lambda}\right) = \mathbb{P}(\text{Poisson}(\lambda) < k) \left(1 - W\frac{k}{\lambda^2}\frac{\mathbb{P}(\text{Poisson}(\lambda) > k)}{\mathbb{P}(\text{Poisson}(\lambda) < k)}\right).$$

To have a simple expression of the competitive ratio, we pick $W = W_{k,\ell}$ such that the above derivative cancels at exactly $\lambda = \ell$. Hence

$$\begin{aligned}
&\frac{d}{d\lambda}\mathbb{E}[\min(\text{Poisson}(\lambda), k)] \left(1 + \frac{W_{k,\ell}}{\lambda}\right) \\
&= \mathbb{P}(\text{Poisson}(\lambda) < k) \left(1 - \frac{\ell^2}{\lambda^2}\frac{\mathbb{P}(\text{Poisson}(\ell) < k)}{\mathbb{P}(\text{Poisson}(\ell) > k)}\frac{\mathbb{P}(\text{Poisson}(\lambda) > k)}{\mathbb{P}(\text{Poisson}(\lambda) < k)}\right)
\end{aligned}$$

It remains to show that this critical point corresponds to a maximum. The computations will be almost identical to [AM21].

For $\lambda < \ell$ we have

$$\begin{aligned}
\frac{\ell^2}{\lambda^2}\frac{\mathbb{P}(\text{Poisson}(\ell) < k)}{\mathbb{P}(\text{Poisson}(\ell) > k)}\frac{\mathbb{P}(\text{Poisson}(\lambda) > k)}{\mathbb{P}(\text{Poisson}(\lambda) < k)} &= \frac{\sum_{j>k}\lambda^{j-2}/j!}{\sum_{j>k}\ell^{j-2}/j!}\cdot\frac{\sum_{j<k}\ell^j/j!}{\sum_{j<k}\lambda^j/j!} \\
&< \left(\frac{\lambda}{\ell}\right)^{k-1}\frac{\sum_{j<k}\ell^j/j!}{\sum_{j<k}\lambda^j/j!} \\
&= \frac{\sum_{j<k}(\frac{\lambda}{\ell})^{k-1-j}\lambda^j/j!}{\sum_{j<k}\lambda^j/j!} \leqslant 1.
\end{aligned}$$

The same can be done for $\lambda > \ell$, which shows that $nq = \ell$ indeed yields the optimal static rule with tie-break for F^*.

In the limit as $n \to \infty$, this implies that

$$\mathrm{CR}^S_{k,\ell} \leqslant \frac{\ell}{k} \cdot \frac{\mathbb{E}[\min(\mathrm{Poisson}(\ell), k)] \left(1 + \frac{W_{k,\ell}}{\ell}\right)}{\ell + W_{k,\ell}} = \frac{\mathbb{E}[\min(\mathrm{Poisson}(\ell), k)]}{k}.$$

This last quantity can then be related to $\gamma_j(\ell)$, as

$$\begin{aligned}
\mathbb{E}[\min(\mathrm{Poisson}(\ell), k)] = \sum_{j=0}^{k-1} \mathbb{P}(\mathrm{Poisson}(\ell) > j) &= k - \sum_{j=0}^{k-1} \mathbb{P}(\mathrm{Poisson}(\ell) \leqslant j) \\
&= \sum_{j=0}^{k-1} \left(1 - \sum_{i=0}^{j} \frac{\ell^i}{i!} e^{-\ell}\right) \\
&= \sum_{j=0}^{k-1} \gamma_{j+1}(\ell) \\
&= \sum_{j=1}^{k} \gamma_j(\ell).
\end{aligned}$$

5.9.15 Proof on finite dimension reduction

In this section, we take care of proving the reduction procedure, for the general (k, ℓ) setting, from a general distribution F to a discrete distribution in $[0, 1]$, with a smaller competitive ratio $\mathrm{CR}_{k,\ell}(n)$. This immediately implies the result of Proposition 5.18 for general (k, ℓ).

Let us first define the technique of balayage.

Definition 5.24. For a random variable X and constants $0 \leqslant a < b < \infty$ we denote by $X_{a:b}$ the random variable which takes the same value as X when $X \notin [a, b]$, takes value a with probability $p_a = \mathbb{E}[(b - X)\mathbf{b}1[X \in [a, b]]]/(b - a)$ and takes value b with probability $p_b = \mathbb{E}[(X - a)\mathbf{b}1[X \in [a, b]]]/(b - a)$.

This new random variable conserves some characteristic of the original one: X and $X_{a,b}$ have the same probability of taking values outside $[a, b]$ thus $\mathbb{E}[X\mathbf{b}1_{[X \notin [a,b]]}] = \mathbb{E}[X_{a,b}\mathbf{b}1_{[X \notin [a,b]]}]$, and by definition of p_a and p_b we have $\mathbb{E}[X\mathbf{b}1_{[X \in [a,b]]}] = \mathbb{E}[X_{a,b}\mathbf{b}1_{[X \in [a,b]]}]$. Both properties imply that $\mathbb{E}[X] = \mathbb{E}[X_{a,b}]$.

We can derive that $\mathrm{OPT}_{\ell,n}$ is increasing with balayage, which generalize the proof of [HK82] for $\ell = 1$.

Lemma 5.25. *For $Y = X_{a:b}$,*

$$\mathbb{E}[\sum_{i\in[\ell]} X_{(i)}] \leqslant \mathbb{E}[\sum_{i\in[\ell]} Y_{(i)}]. \tag{5.39}$$

Proof. We denote by $\mathrm{OPT}_\ell(X_1,\ldots,X_n)$ the function which takes into input the variables $\boldsymbol{x} = (X_1,\ldots,X_n)$ and outputs the sum of the top ℓ variables. Clearly, $\mathbb{E}[\mathrm{OPT}_\ell(\boldsymbol{x})] = \mathrm{OPT}_{\ell,n}$. We first show that for all $i \in [n]$, $\mathbb{E}[\mathrm{OPT}_\ell(\boldsymbol{x})] \leqslant \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i})]$, the statement of the proposition then follows by applying multiple times this inequality.

First, let us remark that $\mathrm{OPT}_\ell(\boldsymbol{x})$ can be rewritten as the value of the following linear (integer) program: the objective is $\mathbf{S}^\top\boldsymbol{x}$ with $S_i \in \{0,1\}$ and $\sum_{i\in[n]} S_i = \ell$. Because the objective is convex, and as the supremum of a family of convex functions, OPT_ℓ is convex in $\boldsymbol{x}$. In particular, for some $i \in [n]$, $\varphi(x) := \mathbb{E}[\mathrm{OPT}_\ell(\boldsymbol{x}) \mid X_i = x]$ is convex in x. By convexity and independence of the X_i, we have that

$$\begin{aligned}
\varphi(x) = \varphi(b\cdot\frac{x-a}{b-a} + a\cdot\frac{b-x}{b-a}) &\leqslant \frac{x-a}{b-a}\varphi(b) + \frac{b-x}{b-a}\varphi(a)\\
&= \frac{x-a}{b-a}\mathbb{E}[\mathrm{OPT}_\ell(b,\boldsymbol{x}_{-i})] + \frac{b-x}{b-a}\mathbb{E}[\mathrm{OPT}_\ell(a,\boldsymbol{x}_{-i})].
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\mathbb{E}[\mathrm{OPT}_\ell(\boldsymbol{x})] &= \mathbb{E}[\mathrm{OPT}_\ell(\boldsymbol{x})\mathbf{b1}_{[X_i\in[a,b]]}] + \mathbb{E}[\mathrm{OPT}_\ell(\boldsymbol{x})\mathbf{b1}_{[X_i\notin[a,b]]}]\\
&= \mathbb{E}[\mathrm{OPT}_\ell(\boldsymbol{x})\mathbf{b1}_{[X_i\in[a,b]]}] + \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i}))\mathbf{b1}_{[X_i\notin[a,b]]}]\\
&= \mathbb{E}[\varphi(X_i)\mathbf{b1}_{[X_i\in[a,b]]}] + \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i}))\mathbf{b1}_{[X_i\notin[a,b]]}]\\
&\leqslant \mathbb{E}\left[\left(\frac{X_i-a}{b-a}\mathbb{E}[\mathrm{OPT}_\ell(b,\boldsymbol{x}_{-i})] + \frac{b-X_i}{b-a}\mathbb{E}[\mathrm{OPT}_\ell(a,\boldsymbol{x}_{-i})]\right)\mathbf{b1}_{[X_i\in[a,b]]}\right]\\
&+ \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i}))\mathbf{b1}_{[X_i\notin[a,b]]}]\\
&= p_b\mathbb{E}[\mathrm{OPT}_\ell(b,\boldsymbol{x}_{-i})] + p_a\mathbb{E}[\mathrm{OPT}_\ell(a,\boldsymbol{x}_{-i})] + \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i}))\mathbf{b1}_{[X_i\notin[a,b]]}]\\
&= \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i}))\mathbf{b1}_{[X_i\in[a,b]]}] + \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i}))\mathbf{b1}_{[X_i\notin[a,b]]}]\\
&= \mathbb{E}[\mathrm{OPT}_\ell(X_{a,b},\boldsymbol{x}_{-i}))].
\end{aligned}$$

□

If the distribution is bounded by some constant v, then we can simply consider the variable X_i/v which is in $[0,1]$, and this does not change the value of the competitive ratio. If the distribution is unbounded, we can do a balayage to infinity which will recover the same property in Lemma 5.25. All the mass above a is put into either a with probability $p_a = \mathbb{E}[(b-X_i)\mathbf{b1}[X_i \geqslant a]]/(b-a)$ and into b with probability $p_b = \mathbb{E}[(a-X_i)\mathbf{b1}[X_i \geqslant a]]/(b-a)$. The only requirement is that $p_a \geqslant 0$ which can be guaranteed for b large enough as X_i was assumed unbounded and therefore $\mathbb{P}(Y \geqslant a) > 0$. See Lemma 2.7 in [HK82] for more details. From now on we consider that the support of F is in $[0, 1]$.

We are now be able to show that for well chosen constants $v_0, \dots, v_m$, we obtain a new distribution such that the value of the optimal algorithm remains the same, while the value of the prophet must be bigger due, hence the competitive ratio smaller, due to Lemma 5.25.

The problem of finding the optimal sequence of stopping times $(\tau_j)_{j\in[k]}$ is directly related to the theory of optimal stopping [Cho+71], and it well known that the Backward Dynamic Programming (BDP) stopping rule is optimal. For k items to select we define the following BDP rule:

Definition 5.26. For $i \in [n]$, let V_i^j be the BDP optimal expected reward of sequentially selecting j among i items, defined by $\mathrm{V}_0^j = 0$, $\mathrm{V}_i^0 = 0$ and by the recurrence relation

$$\mathrm{V}_i^j = \mathbb{E}[(\mathrm{V}_{i-1}^{j-1} + X) \vee \mathrm{V}_{i-1}^j].$$

The BDP stopping rule is to select an item at time i when selecting the j-th item if $X_i + \mathrm{V}_{i-1}^{j-1} \geqslant \mathrm{V}_{i-1}^j$.

This sequence of stopping rules is as mentioned above optimal, and therefore the competitive ratio can be rewritten as the problem of minimizing $(\ell/k)\,\mathrm{V}_n^k(F)/\,\mathrm{OPT}_\ell(F)$.

We will use the following convenient notation for $i \in [n-1]$ and $j \in [k]$:

$$\Delta_i^j := \mathrm{V}_i^j - \mathrm{V}_i^{j-1}\,.$$

For some finite set $B = \{x_i \in \mathbf{b}R, i \in [m]\}$ of m real values , we denote by X_B the successive balayage from left to right of X over B. For instance for $B = \{a, b, c\}$ with $a < b < c$, we have $X_B = (X_{a,b})_{b,c}$.

Remark. It is crucial for the balayage to be done on the ordered values: if we consider $(X_{a,c})_{a,b}$, then $p_b = \mathbb{E}[(X_{a,c} - a)\mathbf{b}1_{[X_{a,c}\in[a,b]]}]/(b-a)$. Because $X_{a,c}$ is already balayed, and $b < c$, $X_{a,b}$ only takes the value a over the interval $[a, b]$, which implies that $p_b = 0$. Whereas for X_B, we can have $p_b \neq 0$. Actually what really matters is for the balayage to be done always with the closest value, but doing it from increasing values gives a proper process to follow.

Lemma 5.27. *For* $B_\Delta = \{0\} \cup \{1\} \cup \{\Delta_i^j, (i,j) \in [n-1] \times [k]\}$*, we have* $\mathrm{V}_n^k(X_{B_\Delta}) = \mathrm{V}_n^k(X)$ *and* $\mathrm{OPT}(X_{B_\Delta}) \geqslant \mathrm{OPT}(X)$*.*

Proof. The second part of the proposition is clear from the fact that OPT is increasing when applying balayage (Lemma 5.25), and that X_B stems from successive balayage. It remains to show the first part.

We denote the ordered elements of B_Δ with $0 = x_1 < x_2 \dots x_{m-1} < x_m = 1$ and consider the sets $B_r = \{x_s, s \in [r]\}$. We show by induction that $\mathrm{V}_n^k(X_{B_r}) = \mathrm{V}_n^k(X)$. The initialization for $\{x_1, x_2\}$ can be proved almost identically to the second induction step, see below. Let us assume that the property is true for $r-1$, with $r \geqslant 3$.

We have $X_{B_r} = (X_{B_{r-1}})_{x_{r-1},x_r}$. Hence, we need to show that for $Y = X_{B_r}$, we have that Y_{x_{r-1},x_r} preserves $\mathrm{V}_n^k(Y)$.

We do a second induction to show that for all $i \in [n]$ the following property is true: for all $j \in [k]$, $\mathrm{V}_i^j(Y) = \mathrm{V}_i^j(Y_{x_{r-1},x_r})$. The initialization is true as $\mathrm{V}_1^j(Y) = \mathbb{E}[Y] = \mathbb{E}[Y_{x_{r-1},x_r}] = \mathrm{V}_1^j(Y_{x_{r-1},x_r})$ where the second inequality comes from balayage preserving expectation. Let us assume that the property is true for $i-1$. We have by the recurrence relation and the induction hypothesis that

$$
\begin{aligned}
&\mathrm{V}_i^j(Y_{x_{r-1},x_r}) = \mathbb{E}[(Y_{x_{r-1},x_r} + \mathrm{V}_{i-1}^{j-1}(Y_{x_{r-1},x_r})) \vee \mathrm{V}_{i-1}^j(Y_{x_{r-1},x_r})] \\
&= \mathbb{E}[(Y_{x_{r-1},x_r} + \mathrm{V}_{i-1}^{j-1}(Y)) \vee \mathrm{V}_{i-1}^j(Y)] \\
&= \mathbb{E}\left[\left((Y_{x_{r-1},x_r} + \mathrm{V}_{i-1}^{j-1}(Y)) \vee \mathrm{V}_{i-1}^j(Y)\right) \mathbf{b1}_{[Y_{x_{r-1},x_r} < \Delta_{i-1}^j]}\right] \\
&+ \mathbb{E}\left[\left((Y_{x_{r-1},x_r} + \mathrm{V}_{i-1}^{j-1}(Y)) \vee \mathrm{V}_{i-1}^j(Y)\right) \mathbf{b1}_{[Y_{x_{r-1},x_r} \geqslant \Delta_{i-1}^j]}\right] \\
&= \mathrm{V}_{i-1}^j(Y)\mathbb{P}(Y_{x_{r-1},x_r} < \Delta_{i-1}^j) + \mathbb{E}[((Y_{x_{r-1},x_r} + \mathrm{V}_{i-1}^j(Y))\mathbf{b1}_{[Y_{x_{r-1},x_r} \geqslant \Delta_{i-1}^j]}] \\
\text{or} \quad &\mathrm{V}_{i-1}^j(Y)\mathbb{P}(Y_{x_{r-1},x_r} \leqslant \Delta_{i-1}^j) + \mathbb{E}[((Y_{x_{r-1},x_r} + \mathrm{V}_{i-1}^j(Y))\mathbf{b1}_{[Y_{x_{r-1},x_r} > \Delta_{i-1}^j]}]
\end{aligned}
$$

If $\Delta_{i-1}^j \notin [x_{r-1}, x_r]$, then we directly have the desired equality. Otherwise $\Delta_{i-1}^j \in [x_{r-1}, x_r]$. In this case because of the construction of B_Δ, we have that Δ_{i-1}^j is either x_r or x_{r-1}.

If $\Delta_{i-1}^j = x_{r-1}$, by the balayage being equal outside $[\Delta_{i-1}^j, x_r]$ we have that $\mathbb{P}(X_{x_{r-1},x_r} < \Delta_{i-1}^j) = \mathbb{P}(X < \Delta_{i-1}^j)$ and $\mathbb{P}(X_{x_{r-1},x_r} \geqslant \Delta_{i-1}^j) = \mathbb{P}(X_{x_{r-1},x_r} \geqslant \Delta_{i-1}^j)$, and in addition with the expectation being equal over $[\Delta_{i-1}^j, x_r]$ we can deduce that $\mathbb{E}[Y_{x_{r-1},x_r}\mathbf{b1}_{[Y_{x_{r-1},x_r} \geqslant \Delta_{i-1}^j]}] = \mathbb{E}[Y\mathbf{b1}_{[Y \geqslant \Delta_{i-1}^j]}]$. Using the first alternate formula described above we obtain $\mathrm{V}_i^j(Y_{x_{r-1},x_r}) = \mathrm{V}_i^j(Y)$. If $\Delta_{i-1}^j = x_r$ we obtain similar properties and use the second alternate formula. In all case we have the equality. We can conclude the second induction, and also conclude the first induction as well. □

Using the above proposition, we know that we can lower the competitive ratio by applying this specific B_Δ balayage on X. All those distributions are supported on the values described by B_Δ. Notice that whenever $j \geqslant i$, then $\Delta_i^j = \Delta_i^i$, so those two values are not distinct. We prove now prove Proposition 5.18.

Proposition 5.28. *The value of* $\mathrm{CR}_{k,\ell}$ *is attained by a discrete distribution with a support of* $2 + k(k-1)/2 + k(n-k)$ *points on* $[0, 1]$.

Proof. This is immediate by applying Lemma 5.27 and Lemma 5.25, and because

$$|B_\Delta| = 2 + \sum_{i \in [n-1]} \sum_{1 \leqslant j \leqslant \max(i,k)} 1 = 2 + k(k-1)/2 + k(n-k). \qquad \square$$

It is possible that other reductions are more efficient in terms of numbers of values.

Hence we can consider an optimization problem over $2(2 + k(k-1)/2 + k(n-k))$ parameters instead (to take into account different possible values with different associated distributions).

Interestingly, the gaps respect some monotonicity property:

Proposition 5.29. *The* Δ_i^j *are increasing in* i *and decreasing in* j.

Let us first show that it is increasing in j. We have that $\Delta_i^j - \Delta_i^{j-1} = \mathrm{V}_i^j - 2\,\mathrm{V}_i^{j-1} + \mathrm{V}_i^{j-2}$. Let us compare $\mathrm{V}_i^j + \mathrm{V}_i^{j-2}$ to $2\,\mathrm{V}_i^{j-1}$. The first quantity correspond to the supremum of stop rules, where it is allowed to select 2 times the same item for the $j-2$ first items, and then is allowed to select 2 more items at different times. The second quantity correspond to stop rules allowed to select 2 times the same item for the first $j-1$ items encountered. This is strictly more lax in terms of constraints compared to the first quantity, hence we have that Δ_i^j is decreasing in j.

We now show that it is increasing in i.

$$\begin{aligned}
\Delta_{i-1}^j &= \mathrm{V}_{i-1}^j - \mathrm{V}_{i-1}^{j-1} = \mathbb{E}[(X + \mathrm{V}_i^{j-1}) \vee \mathrm{V}_i^j] - \mathbb{E}[(X + \mathrm{V}_i^{j-2}) \vee \mathrm{V}_i^{j-1}] \\
&= \mathrm{V}_i^j - \mathrm{V}_i^{j-1} + \mathbb{E}[(X + \mathrm{V}_i^{j-1} - \mathrm{V}_i^j)_+ - (X + \mathrm{V}_i^{j-2} - \mathrm{V}_i^{j-1})_+] \\
&= \Delta_i^j + \mathbb{E}[(X + \mathrm{V}_i^{j-1} - \mathrm{V}_i^j)_+ - (X + \mathrm{V}_i^{j-2} - \mathrm{V}_i^{j-1})_+].
\end{aligned}$$

Using that if $z > y$, then $z_+ > y_+$, and because

$$X + \mathrm{V}_i^{j-1} - \mathrm{V}_i^{j} - X - \mathrm{V}_i^{j-2} + \mathrm{V}_i^{j-1} = \Delta_i^{j-1} - \Delta_i^{j} \geqslant 0,$$

we can conclude. The inequality comes from the monotonicity of Δ_i^j in j.

5.10 Further related works

While the i.i.d. version of the prophet inequality has received significant attention, other variants have been studied extensively. If $1/2$ is the best competitive ratio when the values are not distributed identically and arrive in a fixed sequence, [Esf+17; Ehs+17] show that when the X_i are presented in a random order, named prophet secretary problem, a competitive ratio of at least $1 - 1/e$ can be achieved. [Cha+10; Siv+21] study the free order prophet where the order of arrival of the X_i can be freely chosen. Recently, [BC22; GMS23] have shown that both of these variants are intrinsically different, in that their worst-case competitive ratio are distinct. An important remaining question, is whether the free order variant is as hard as the i.i.d. case. This is related to our work, as any upper bound on the i.i.d. case directly translates into an upper bound on the free order prophet.

In an orthogonal direction, it is possible to examine prophet settings with increasingly complex combinatorial constraints or payoffs. There has been a rich stream of literature on the multi-unit prophet, which assumes that the decision maker and the prophet both actually have a budget of $k \in \mathbf{b}N$ items, which was initiated by [HKS07]. Lower bounds for the competitive ratio explicit in k of order $1 - O(1/\sqrt{k})$ were subsequently given by [Ala11] for an adaptive algorithm, and [CDL20] then proved that $1 - O(\log(k)/\sqrt{k})$ can be reached using only a single threshold. More recently [JMZ21] gave tight constants that are solutions to a limiting ODE. [JMZ22] also proposes optimization problems that compute the competitive ratio for any k but only for a given n. Different types of constraints are also studied such as [KW12a] which assumes that the allocation must respect matroid constraints, or [CC23] who proved competitive ratio guarantees for an online combinatorial auction.

The idea of considering weaker benchmarks, as proposed by [Ken85] and our paper, can be readily considered for any of these different combinatorial or distributional assumptions. The more general framework where the decision maker and the prophet can respectively select k and ℓ items was introduced by [Ken87] in the non i.i.d. case, but significant results were only proven for $\ell = 1$. This is of the same flavor as the (J, K)-secretary problem introduced by [BJS10], where the goal is to find an element in the top K with only J tries. Selecting one item among the top k has also been studied for the prophet setting in [Esf+19], where the probability of getting an item from the top k is of order $1 - O(\exp(-k))$. A recent work by [Har25] proposes a new sharding technique to obtain better prophet inequalities, and in particular achieves a $O(1 - k^{k/5})$ lower bound in the i.i.d. setting for the problem where the decision maker recovers as a value the maximum of k selected items. Another setting where a rate of $1 - e^{-k}$ is achievable is in [Ala+22], where k different thresholds are fixed in advance, and the decision maker is allowed to make k passes over the data.

Finally, we mention that there has been a recent nice concurrent work by [BPV24] on the k multi-unit i.i.d. prophet who, using a complementary approach through a linear program characterization of $\mathrm{CR}_{k,k}(n)$, achieve similar results. They obtain the same limit system of ODE for $k = \ell$, and additionally provide an estimate of the error between the asymptotic value $\liminf_n \mathrm{CR}_{k,k}(n)$ and $\mathrm{CR}_{k,k}(n)$. They also leverage this result to prove a tight approximation ratio for the stochastic sequential assignment problem.

There has also been a lot of focus [AKW14; Cor+19b] on sample prophet inequalities, where decision makers do not have access to the distribution themselves, but only samples of the distribution. A remarkable result from [RWW20] is that a single sample per distribution is enough to achieve the $1/2$ competitive ratio in the original prophet setting. They also show how to use the quantile strategies from [Cor+17] to obtain sample prophet inequalities in the i.i.d. case. This is especially relevant for this work, as the strategies we propose are also quantile algorithms, and therefore the proof from [RWW20] can likely be extended by using Algorithm 10.

Acknowledgements

This work has been partially supported by MIAI @ Grenoble Alpes (ANR-19-P3IA-0003), by the Hi! PARIS Center, by the French National Research Agency (ANR) through grant ANR-20-CE23-0007, ANR-19-CE23-0015, ANR-23-CE23-0002, ANR-19-CE23-0026 and PEPR IA FOUNDRY project (ANR-23-PEIA-0003). This work has also been funded by the European Union through the ERC OCEAN 101071601 grant. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council

Executive Agency. Neither the European Union nor the granting authority can be held responsible.

Equitable Auctions

6

Abstract. We initiate the study of how auction design affects the division of surplus among bidders. We propose a parsimonious measure for equity and apply it to standard auctions for homogeneous goods. Our surplus-equitable mechanism is efficient, Bayesian-Nash incentive compatible, and achieves surplus parity among winners ex-post. The uniform-price auction is equity-optimal if and only if bidders have a common value. Against intuition, the pay-as-bid auction is not always equity-preferred if bidders have private values. In auctions with price mixing between pay-as-bid and uniform prices, we provide prior-free bounds on the equity-preferred pricing under a common regularity condition on signals.

Contents

6.1 Introduction

Equity concerns in auctions chave increasingly entered policy debates and are of critical importance to participation and stability in downstream markets. In practice, auctions are used, for example, to sell government debt, electricity, emission permits, oil, timber, coffee, art, and production input factors. Although auction design is a cornerstone of economics research, the *division of surplus between bidders* has not been studied. Generally, auctions are held to elicit agents' values or costs for buying or selling goods. However, the auction mechanism, consisting of an allocation and pricing rule, can result in an asymmetric distribution of welfare, even among winning bidders: for example, in single-price auctions, high-value bidders obtain a larger surplus (i.e., value minus price) than low-value bidders. This points to possibly unintended implications for the welfare distribution in the auction. Nonetheless, we show that auctions can be made equitable by design.

This article initiates the study of surplus distribution between bidders in multi-unit auctions. We focus on the class of standard auctions with independent signals, which are revenue equivalent and efficient. Thus, designing the equity objective is costless in terms of potential trade-offs. We propose a family of equity metrics that are based on parsimonious, pairwise comparisons of realized surpluses (utilities). First, we characterize the direct surplus-equitable mechanism. This incentive-compatible mechanism uses windfall subsidies to achieve ex-post identical surpluses among winners of any type realization. Second, we turn to uniform and pay-as-bid auctions and combinations of these, that is, mixed-price auctions. In this class, we derive prior-free results on equity-preferred pricing, with strong policy implications for multi-unit auctions used in practice.

Uniform pricing and pay-as-bid (discriminatory) pricing are the prevalent multi-unit auction formats.[1] In the uniform-price auction, all winners pay the first rejected bid, and in the pay-as-bid auction, all winners pay the price they bid. The question of which pricing rule leads to more efficient power market outcomes, less collusion, and more revenue has been debated for decades (e.g., [Kah+01; Aus+14]), but fairness and redistribution concerns have received little attention, albeit a central policy concern. For example, the Small Business Act in the US requires the government to award 23% of procurement contracts each year to small businesses that are socially and economically disadvantaged or minority-owned [PV12; US 24]. In spectrum auctions, allocative fairness in the distribution of licenses is particularly

[1] In treasury auctions both designs are common [OEC21], while most electricity markets feature the uniform pricing rule. Further examples include auctions for emission certificates, or online advertisement.

important, as it affects competition in the downstream market [GSM21; Kas23]. Similarly, the distribution of surplus can influence market stability and competition in post-auction markets: bidders disadvantaged in the surplus distribution may face higher borrowing costs, especially in inefficient capital markets. This is aggravated in auctions that occur only once, as a low surplus can be detrimental to a company's survival in the post-auction market.

Generally, redistribution after the auction, withstanding legal feasibility, may distort bidding incentives and efficiency. However, the taxation of rents infra-marginal bidders earn in electricity auctions was debated, for example, in the context of recovering infrastructure investment costs [New14; RPD17].

The proposed tax on companies' observable surplus is equivalent to a mixed auction, in which prices are set by a combination of uniform and pay-as-bid pricing, a design suggested for some electricity markets by [HT23]. This class of mechanisms is one of the focal points of this article.

Our setup is as follows. We consider standard and winners-pay multi-unit auctions (so-called k-unit auctions) for the sale of indivisible, identical goods with a composition of private and common values. Each bidder has unit demand and receives a private and independent signal drawn from a publicly known distribution.[2]

A bidder's value linearly interpolates between the extremes of pure private and pure common value, that is, between their private signal and the average signal in the market.[3] We study direct incentive-compatible mechanisms and the class of mixed k-unit auctions. Mixed auctions combine uniform and pay-as-bid auctions and incorporate those as special cases. For a given δ, we call the convex combination of uniform and pay-as-bid pricing *δ-mixed pricing*, and the corresponding auction *δ-mixed auction*. The parameter δ describes the degree of price discrimination: $\delta = 0$ corresponds to uniform pricing and $\delta = 1$ to pay-as-bid pricing. In mixed auctions, we study the symmetric Bayesian equilibrium, which is found to be unique.

All considered auctions, under classical assumptions, achieve the same expected revenue [MS02] and allocate items to the highest-value bidders; thus, there are no trade-offs with revenue or efficiency.

[2] We provide a discussion of the unit-demand assumption in section 6.6.4.

[3] This model can represent common resale opportunities. For example, an emission certificate is valuable for a company's production process (private value), but it can also be resold after the auction, with the resale value being common to the market. The resale model appears in previous work, for example, by [BR91; Kle98; BK02; GO03].

Contributions Our first conceptual contribution is the introduction of a new, parsimonious measure of equity, *dominance in pairwise differences*: a mechanism A dominates another mechanism B in pairwise differences, if, in equilibrium, all absolute pairwise differences in ex-post utilities of winning bidders are weakly smaller in A than in B (definition 6.6), with one strict comparison. We say that A is *equity-preferred* to B. Our results hold for the family of anonymous equity metrics that are constructed by aggregation of pairwise differences with any increasing function. This family includes prominent inequality measures, e.g., the empirical variance, the Gini index or comparison of the top and bottom deciles (see also [Lor05; Gin12b; Gin21; Pig12b; Dal20b; Atk70; SF73]). As an example of an aggregator, we study the *winners' empirical variance (WEV)* of surplus, which satisfies monotonicity in transfers from richer to poorer agents (proposition 6.8) (Pigou-Dalton principle, cf. [Mou04]).

Our main results are as follows.

1. We characterize the direct and Bayes-Nash incentive-compatible mechanism that distributes realized surpluses equally among the winning bidders (theorem 6.11). The surplus-equitable mechanism allocates the items to the highest bidders and charges them a payment consisting of three components: firstly, each bidder pays their private value, thus equalizing ex-post utilities; secondly, a uniform payment that cancels out the idiosyncratic payment in expectation; and finally, the expected value corresponding to the first rejected bid, akin to a "second-price" payment. The key and surprising insight is that the uniform payments cancel out the idiosyncratic payment part, *for any given signal*, in expectation.

2. We prove that the uniform-price auction is equity-preferred (is dominant in pairwise differences) if and only if the bidders' values are pure common value (theorem 6.13). In this case, the surplus-equitable mechanism is equivalent to the first-rejected-bid uniform-price auction, which, as any other uniform-price auction, equalizes bidders' realized surpluses. By contrast, for pure private values, the pay-as-bid auction is not generally equity-preferred in the class of mixed auctions (proposition 6.16).

3. We then provide prior-free results in the class of log-concave distributions. Given any proportion of private values $(1-c)$, the $(1-c)$-mixed auction is equity-preferred over any mixed auction with less than a $(1-c)$ share of price discrimination (theorem 6.18). That is, the pricing for goods with a higher proportion of private value should contain more price discrimination in order to achieve an equitable distribution of realized surpluses. In particular,

for pure private values, if signals are drawn from a log-concave distribution (corollary 6.19) the pay-as-bid auction is equity-preferred to all mixed-priced auctions. Any level of price discrimination up to $2(1-c)$ is equity-preferred to uniform pricing.

Finally, we investigate equity in terms of winners' empirical variance (WEV) in numerical experiments. For a variety of signal distributions and common value proportions, we compute the landscape of WEV-minimal mixed pricing, which can be seen to be unique.

6.1.1 Related literature

This article relates and contributes to several strands of existing work, including a recent literature on redistributive market design, the study of fairness concerns and allocative equity in auctions, and of fair allocations more generally. Further, we contribute to the mechanism design literature on ex-post payment design and the analysis of uniform, pay-as-bid, and mixed-price auctions.

Broadly, our contribution fits into a recent strand of the economic literature on redistributive concerns in market design. In this literature, the focus is often on efficiency and equity trade-offs; e.g., in a large buyer-seller market for a single object, with agents differing in their marginal utilities of money (and values), [DKA21] characterize the optimal efficiency-equity trade-off, and [ADK24], characterize when non-market mechanisms, as opposed to market-clearing prices, are optimal for a designer to allocate a fixed supply (also in a large market). Such non-market mechanisms forgo efficiency for the sake of improving equity. Our approach differs in that we focus on a class of efficient mechanisms in a small market with a finite number of bidders and demonstrate how to improve equity, up to achieving perfect surplus parity.

Fairness concerns in auctions have been addressed through design instruments such as subsidies and set-asides. In a model with explicit target group favoritism, [PV12] show that the optimal mechanism is a flat or a type-dependent subsidy, depending on the precise nature of the favoritism constraint. [ACL13] come to similar conclusions in an empirical study of US Forest Service timer auctions, where set-asides for small bidders would reduce efficiency and revenue, while subsidies would increase revenue and profits of small bidders with little detriment to efficiency. In contrast with this literature, we focus on equity among winners and show that a carefully designed pricing rule can achieve a more equitable distribution of surplus.

The literature on fair allocation has introduced many concepts of fairness, including, for example, *envy-freeness*, *equal division*, or *no domination* (for a survey, see [Tho11]). Our notion of fairness is orthogonal to envy-freeness, which requires that no agent prefers another agent's allocation (object and price). In our market, uniform pricing is the unique pricing scheme that results in envy-freeness among winners, and with pure private values, it results in envy-freeness among all participants.[4] However, envy-freeness does not take into account that bidders may have different signal (value) realizations. Subscribing to the notion of envy-freeness, we would accept that realized utilities may be very unequal, depending on the realization of the private value component. In contrast to this view, we consider the realization of utilities as the baseline for fairness considerations, which relates to *equal treatment of equals* [cf., e.g., Tho11] in an ex-post view.

The design of our ex-post surplus-equalizing payment rule also contribute to a strand of the mechanism design literature that developed important, nuanced ex-post implementations of truthful mechanisms. In their seminal article, [AG79] show that ex-post budget balance can be achieved in a direct truthful mechanism. [EF99] prove that for every IC mechanism there exists a mechanism which provides deterministically the same revenue.

Outline The remainder of the article is organized as follows. In section 6.2, we introduce the model and derive equilibrium bidding strategies in mixed auctions. We introduce our notion of surplus equity in section 6.3 and develop the surplus-equitable mechanism in section 6.4. In section 6.5, we describe our prior-free results for uniform, pay-as-bid, and mixed auctions and prove them in section 6.7. section 6.6 provides a discussion and section 6.6.5 concludes.

6.2 Setup

6.2.1 Model

A finite number of bidders $[n] := \{1, \ldots, n\}$ compete for a fixed supply of items $[k] := \{1, ..., k\}$, where $2 \leqslant k < n$. Each bidder only demands one item. Bidder i receives a private signal s_i, which is drawn independently from a positive and bounded or unbounded support; denote its upper limit by $\overline{s}$. Signals are iid with

[4] With a proportion of common value, depending on the realization of signals, winners may experience the winners' curse and prefer not to have won an item.

an absolutely continuous probability distribution F with density f. We call $(0, \overline{s})$, i.e., all signals s so that $0 < F(s) < 1$, the open support of F, and assume that $f > 0$ over $(0, \overline{s})$. We also assume that the signals have a finite second moment $\mathbb{E}[s^2] < \infty$.

For $\mathbf{s} := \{s_i\}_{i \in [n]}$, a collection of iid signals, we denote by $Y_m(\mathbf{s})$ the m-th highest value of the collection $\mathbf{s}$ (with n entries). For example, $Y_1(\mathbf{s})$ is the maximum and $Y_n(\mathbf{s})$ is the minimum of the collection $\mathbf{s}$. Note that $Y_m(\mathbf{s})$ is a random variable, and we denote its probability distribution $G_m^n(y)$ with corresponding density $g_m^n(y)$. G_m^n is given by

$$G_m^n(y) = \sum_{j=0}^{m-1} \binom{n}{j} F(y)^{n-j}(1 - F(y))^j. \tag{6.1}$$

where each summand is the probability that j signals are above y. An expression for $g_m^n(y)$ is given in section 6.9.6. The value of bidder i for an item is given by the valuation function $v(s_i, \mathbf{s}_{-i})$, where $\mathbf{s}_{-i} := (s_j)_{j \neq i}$, and $v(s_i, \mathbf{s}_{-i})$ is symmetric in other bidders' signals $\mathbf{s}_{-i}$.

Assumption 6.1. Values $v(s_i, \mathbf{s}_{-i})$ are given by $v(s_i, \mathbf{s}_{-i}) = (1 - c)s_i + \frac{c}{n} \sum_{j \in [n]} s_j$, where $c \in [0, 1]$ is the *proportion of the common value*.

Our model interpolates between a common value and private values, where the proportion of the common value c encodes to what extent the value of any given bidder is influenced by the signals of the other bidders. In particular, $c = 1$ defines a pure common value and $c = 0$ pure private values. We note that the value function satisfies the *single-crossing* condition as for all $i, j \in [n]$, $i \neq j$, and for all $\mathbf{s}$, $\partial v(s_i, \mathbf{s}_{-i})/\partial s_i \geqslant \partial v(s_j, \mathbf{s}_{-j})/\partial s_i$.

Auction mechanisms. Auction mechanisms are represented by allocations and transfers $\{\pi_i(s_i, \mathbf{s}_{-i}), p_i(s_i, \mathbf{s}_{-i})\}_{i \in [n]}$, where $\pi_i(s_i, \mathbf{s}_{-i})$ is defined as the probability that an item is allocated to the bidder i when the reported signals are s_i and $\mathbf{s}_{-i}$, and $p_i(s_i, \mathbf{s}_{-i})$ is the corresponding price charged to the bidder. The price is symmetric in $\mathbf{s}_{-i}$, i.e., $p_i(s_i, \mathbf{s}_{-i}) = p(s_i, \mathbf{s}'_{-i})$ for all permutations $\mathbf{s}'_{-i}$ of $\mathbf{s}_{-i}$. We require that auction mechanisms be standard and winners-pay. An auction is *standard* if the k highest bids win the items and *winners-pay* if only the winners pay (at most their bid). Any standard auction, in any symmetric and increasing equilibrium and values satisfying the single-crossing condition, is *efficient* [Kri09], i.e., the bidders with the k highest values $v(s_i, \mathbf{s}_{-i})$ are assigned the items.

We consider two classes of mechanisms. First, we consider truthful, direct mechanisms in which bidders submit their signal. Second, we consider k-unit *mixed auctions* (defined below) in which each bidder submits a bid b_i, resulting in the vector of submitted bids $\mathbf{b}$. Restricting our attention to symmetric and monotonically increasing bidding strategies $b_i = b(s_i)$, we can write allocations and prices in both classes of mechanism as functions of signals only. The allocation of bidder i is given by $\pi_i(s_i, \mathbf{s}_{-i}) = \mathbb{1}\{s_i > Y_k(\mathbf{s}_{-i})\}$ when signals s_i and $\mathbf{s}_{-i}$ are reported. A bidder's utility (or surplus) when reporting signal $\widehat{s}_i$, and the remaining $n-1$ bidders reporting signals $\mathbf{s}_{-i}$, is given by

$$u_i(s_i, \widehat{s}_i, \mathbf{s}_{-i}) = \mathbb{1}\{\widehat{s}_i > Y_k(\mathbf{s}_{-i})\} \cdot v(s_i, \mathbf{s}_{-i}) - p_i(\widehat{s}_i, \mathbf{s}_{-i}). \tag{6.2}$$

Given a signal s_i, recall that we denote by $Y_k(\mathbf{s}_{-i})$ the k-th highest among the signals $\mathbf{s}_{-i}$. $Y_k(\mathbf{s}_{-i})$ has probability distribution G_k^{n-1} and density g_k^{n-1}.

Furthermore, we denote equilibrium bidding strategies by $(\beta_i)_{i\in[n]} = \boldsymbol{\beta}$.

Definition 6.2 (Mixed auctions)**.** In the k-unit δ-*mixed auction*, parameterized by a given $\delta \in [0,1]$, each bidder i pays $p_i(\mathbf{b}) = (\delta b_i + (1-\delta)Y_{k+1}(\mathbf{b}))\,\mathbb{1}\{b_i > Y_{k+1}(\mathbf{b})\}$.

At one boundary, for $\delta = 0$, this resolves to *first-rejected-bid uniform pricing* or short *uniform pricing*, where each winning bidder i pays the $(k+1)$-th highest bid $Y_{k+1}(\mathbf{b})$. At the other boundary, for $\delta = 1$, this resolves to *pay-as-bid pricing*, where each winning bidder i pays their bid b_i. Finally, if $\delta \in (0,1)$, we say that the auction and the pricing are *strictly mixed*.[5]

6.2.2 Interim values, payments, and utilities

For all $x, y \in [0,1]$, we define the expected value given $s_i = x$ and $Y_k(\mathbf{s}_{-i}) = y$ as follows:

$$\widetilde{V}(x,y) := \mathbb{E}_{\mathbf{s}}[v(s_i, \mathbf{s}_{-i}) \mid s_i = x, Y_k(\mathbf{s}_{-i}) = y]. \tag{6.3}$$

The expected value is taken over $n-2$ signals not including the bidder's own signal and the k-th highest among their $n-1$ opponents. Observe that because $v(s_i, \mathbf{s}_{-i})$ is continuous and non-decreasing, $\widetilde{V}(x,y)$ is continuous and non-decreasing in x and y.[6] We define $V(y) := \widetilde{V}(y,y)$, the expectation of the value of an item conditional on the bidder winning against the relevant competing signal, the k-th highest among its

[5] Mixed-price auctions, originating from [WZ02] and [VW02], have also appeared in a series of articles modeling a divisible and stochastic supply [RPD17; MTL20; Woo21].

[6] In fact, it is strictly increasing in x.

competitors. Furthermore, we introduce *interim payments* $P_i(s_i) = \mathbb{E}_{\mathbf{s}_{-i}}[p_i(s_i, \mathbf{s}_{-i})]$ and *interim utilities* $U_i(s_i, \widehat{s}_i) = \mathbb{E}_{\mathbf{s}_{-i}}[u_i(s_i, \widehat{s}_i, \mathbf{s}_{-i})]$ with $U_i(x)$ being the shorthand of $U_i(x, x)$.

Interim incentive compatibility (IC) requires $U_i(s_i, s_i) \geqslant U_i(s_i, \widehat{s}_i)$ for all $s_i, \widehat{s}_i$, and interim individual rationality (IR) demands $U_i(s_i, s_i) \geqslant 0$ for all s_i. It is standard from an application of the envelope theorem [MS02] that the auctions we consider result in the same interim payment (and interim surplus) fixing a bidder's signal (see Section 6.8).[7]

6.2.3 Equilibrium bidding

We derive the unique Bayes-Nash equilibrium in increasing and symmetric bidding strategies in δ-mixed auctions. This equilibrium is the center of our analysis of surplus equity in mixed auctions in section 6.5.

Proposition 6.3 (e.g., [Kri09]). *The unique equilibrium bidding strategy in the uniform-price auction, i.e., the case $\delta = 0$, is given by $\beta^U(s) := \widetilde{V}(s, s) = \mathbb{E}[v(s_i, \mathbf{s}_{-i}) \mid s_i = s, Y_k(\mathbf{s}_{-i}) = s]$.*

Note that the equilibrium is unique in the class of increasing and symmetric strategies and weakly dominant with pure private values [Kri09]. The proof of the next proposition is given in section 6.9.2.

Proposition 6.4. *The unique symmetric equilibrium bidding strategy in the δ-mixed auction, for $\delta \in (0, 1]$, is given by*

$$\beta^\delta(s) = V(s) - \frac{\int_0^s V'(y) G_k^{n-1}(y)^{\frac{1}{\delta}} \, \mathrm{d}y}{G_k^{n-1}(s)^{\frac{1}{\delta}}}. \tag{6.4}$$

Note that β^δ converges to β^U as $\delta \to 0$. We illustrate it in the following example.

Example 6.5 (label=uniform-example). We consider the simplest, non-degenerate setting with $n = 3$ bidders competing for $k = 2$ items. The bidders' signals are distributed uniformly on the support $[0, 1]$. For the uniform-price auction, one can easily compute $\beta^0(s) = (1 - c)s + \frac{c}{6}(5s + 1)$ and $\beta^1(s) = (1 - \frac{c}{6}) \frac{s - \frac{2}{3}s^2}{2 - s} + \frac{c}{6}$. Note

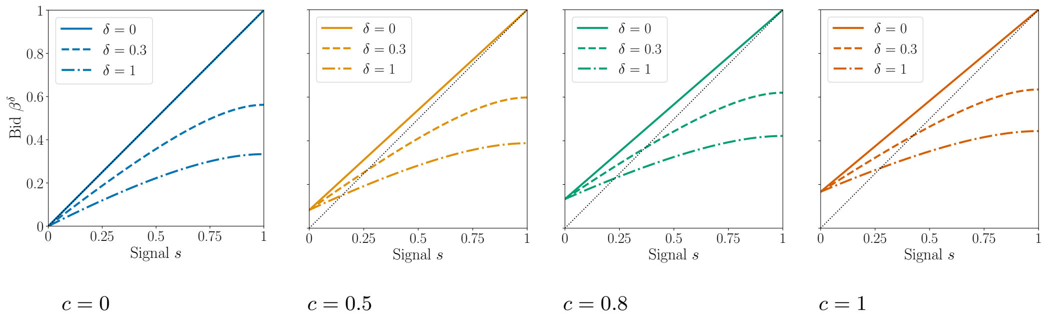


Fig. 6.1: Equilibrium bid functions, β^δ, for uniform signal distributions as a function of the signal, s, for common value parameters $c \in \{0, 0.5, 0.8, 1\}$.

that β^0 is linear due to uniform signals. fig. 6.1 illustrates the bid functions for four different values of c.

In the case of pure private values, bidding truthfully is a dominant strategy in the uniform-price auction. Increasing price discrimination, δ, decreases the bid corresponding to a given signal below the private value ("bid shading"). The extent of bid shading increases in δ. A higher common value component shifts equilibrium bids for low signal realizations above the diagonal, because a low-signal bidder's expectation of the average signal is higher than their private signal. For $s \to 1$, however, the expected value of an item conditional on tying with the price-setting competing signal converges to s, i.e., truthful bidding. Note that the latter need not be true for different n and k.

With a pure common value, the winner's curse becomes especially apparent. In equilibrium, bidders are attempting to salvage the winner's curse but cannot escape it. In fact, with a pure common value, winners' ex-post utilities are *decreasing in signals* for any $\delta > 0$.

6.3 Surplus equity

We propose an equity notion that we call *dominance in (absolute) pairwise differences*, or short *pairwise differences*. Our results hold for the family of equity measures that are defined by any increasing function of pairwise differences in ex-post surpluses. We call the collection $\{u_i(\mathbf{s})\}_{i\in[n]}$ of ex-post utilities an *outcome*, where each utility depends on the collection of signals $\mathbf{s}$.

[7]This is also shown differently in [Kri09] for the single-unit auction. Note that in settings where signals are affiliated revenue equivalence fails [Kri09, Chapter 6.5].

Definition 6.6 (Dominance in pairwise differences among winners)**.** An outcome $\{u_i(\mathbf{s})\}_{i\in[n]}$ dominates another outcome $\{u'_i(\mathbf{s})\}_{i\in[n]}$ in pairwise differences iff, for all winning signals s_i, s_j with opponents' signals $\mathbf{s}_{-i}, \mathbf{s}_{-j}$, $i, j \in [n]$, it holds that $|u_i(s_i, \mathbf{s}_{-i}) - u_j(s_j, \mathbf{s}_{-j})| \leqslant |u'_i(s_i, \mathbf{s}_{-i}) - u'_j(s_j, \mathbf{s}_{-j})|$, almost surely and with one inequality strict.

Naturally, dominance pairwise differences can be defined similarly based on pairwise comparisons among all bidders. To avoid issues with ties, we consider that dominance in pairwise differences holds as long as it holds almost surely. Furthermore, we say that, for a family of parameterized outcomes $\{u_i^\delta\}_{i\in[n]}$, $\delta \in \Delta$, δ^* is dominant in pairwise differences if u^{δ^*} dominates all outcomes u^δ, $\delta \neq \delta^*$, $\delta \in \Delta$. Pairwise differences induces a partial dominance ranking over outcomes and therefore a dominant δ^* may not always exist.

Several prominent equity axioms (cf., e.g., [PP19]) hold for pairwise differences. First, we note that anonymity is maintained. Any reordering of individuals in the population $[n]$ has no consequence, as pairwise comparisons must hold for any two bidders.[8] The Pigou-Dalton transfer principle requires that any transfer from a wealthier agent to a poorer one must reduce inequality, provided the original welfare ranking between the two agents is maintained, that is, the wealthier agent does not become poorer than the previously poorer agent after the transfer (cf., e.g., [Mou04]). Since dominance in pairwise differences does not establish a complete order of outcomes, a Pigou-Dalton transfer may result in a decrease in some pairwise differences while others increase.

However, our results allow the classification of δ-mixed pricing rules based on pairwise differences and *any increasing function* of pairwise differences.[9] For example, the top decile of realized utilities can be compared to the lowest or the bottom decile of realized utilities, and classic inequity measures such as the Gini index can be constructed. Larger differences can receive a higher weight than smaller ones, e.g., by squaring each pairwise difference.

To exemplify the aggregation of pairwise differences, we focus on the *expected empirical variance* of surplus between the winners, or *winners' empirical variance*

[8] We note that replication invariance and mean independence are not relevant in our setup, as we keep the population size (number of bidders) as well as the endowments (value distributions) fixed.

[9] A related aggregation is used by [FK74], who aggregate positive pairwise differences for a measure of envy per player. Contrasting our measure, [FK74] consider pairwise differences of hypothetical (if an agent had received another agent's bundle) and realized utilities.

(WEV) for short.[10] This metric is defined in expectation, ensuring that it provides a ranking of auction designs for any signal realizations.

Definition 6.7 (Winners' empirical variance)**.**

$$\text{WEV} = E_s\left[\frac{1}{2k(k-1)}\sum_{i=1}^{k}\sum_{j=1}^{k}(u_i(\mathbf{s})-u_j(\mathbf{s}))^2\middle|s_1,\ldots,s_k > Y_{k+1}(\mathbf{s})\right]. \tag{6.5}$$

In addition to being a natural and well-known metric, this aggregation is attractive for two reasons: First, it ensures compliance with the Pigou-Dalton transfer principle, and second, the empirical variance is linked to surplus variance and correlation of surpluses among bidders.

Proposition 6.8. *The* winners' empirical variance *satisfies the* Pigou-Dalton principle*.*

The proof is given in section 6.9.1. In expectation, equilibrium surplus varies due to a bidder's own and the competitors' signals, and surplus between winners may be correlated. As we consider efficient auctions, surplus only varies among the winning bidders. Among those, WEV measures *within-bidder variation* and *across-bidder correlation* of surpluses (see lemma 6.9 below). The analysis of within-agent variation addresses a bidder's individual risk-attitude and goes back to [Vic61].[11] In contrast, an equity measure must take into account the correlation of surpluses between bidders.

Lemma 6.9. *An equivalent expression for the winners' empirical variance is given by* WEV $= \text{Var}[u_1|1\ \textit{wins}] - \text{Cov}[u_1, u_2|1\ \textit{and}\ 2\ \textit{win}]$*.*

The proof is in section 6.9.1. In contrast, the ex-ante variance, $\text{Var}_{\mathbf{s}}[u_i(\mathbf{s})]$, measures surplus variation *within* a given bidder, and is more adequate to measure risk, e.g., across a series of identical, repeated auctions, in which a given bidder redraws their signal in every auction. With pure private values and thus ex-post individual rationality, rankings of auction formats in terms of ex-ante variance or winners' ex-ante variance are identical. Rankings with respect to the empirical variance, however, may differ depending on if only winners are considered, or all bidders. A formal lemma and proof are given in section 6.9.1.

[10] The empirical variance among all bidders (thus including losers) in the auction is EV $= E_{\mathbf{s}}[\frac{1}{n(n-1)}\sum_{i=1}^{n}(u_i(\mathbf{s}) - u_j(\mathbf{s}))^2]$.

[11] In the appendix of his famous article, Vickrey showed that, in a *single-unit auction*, the ex ante variance of surplus is lower under the first-price than the second-price rule, given uniform distributions of private values. See also [Kri09].

Example 6.10 (continues=uniform-example)**.** Revisiting the example with uniformly distributed signals, $n = 3$ bidders, and $k = 2$ items, we compute WEV numerically and illustrate it for different values of the common-value proportion c (fig. 6.2). For pure private and intermediate common values, the pay-as-bid auction

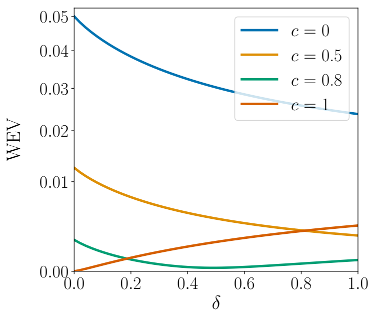


Fig. 6.2: WEV as a function of δ for uniform signals and various common value proportions c

($\delta = 1$) minimizes WEV. For $c = 0.8$, we observe an interior optimum, and for a pure common value, uniform pricing ($\delta = 0$) minimizes WEV. Note that WEV is not necessarily convex in δ for a fixed c.

As the considered mechanisms are revenue equivalent and efficient (see also section 6.8), we can focus on the surplus distribution among bidders without considering potential trade-offs.

6.4 The surplus-equitable mechanism

In the class of efficient auctions, it is possible to distribute surplus among the winning bidders equitably. This means pairwise differences in ex-post utilities are zero, as is any aggregate metric such as the winners' empirical variance. We let $y := Y_k(\mathbf{s}_{-i})$, $G := G_k^{n-1}$, and $g := g_k^{n-1}$.

Theorem 6.11. *In the class of standard k-unit auctions, there exists an incentive-compatible direct mechanism that distributes surpluses equitably among winners. Losers pay nothing, and the corresponding payments are given by*

$$\tilde{p}_i(s_i, \mathbf{s}_{-i}) = \left((1-c)\left(s_i - y - \frac{G(y)}{g(y)}\right) + V(y)\right) \mathbb{1}\left\{s_i > y\right\}. \tag{6.6}$$

The ex-post payment in eq. (6.6) is constructed in a simple but powerful way. First, the term $(1-c)s_i$ removes the idiosyncratic part of each bidder's realized value due to their own signal. To align interim incentives, this term is adjusted by a uniform subsidy $(1-c)(y + G(y)/g(y))$, which cancels the idiosyncratic payment in expectation. Thus, the second-price payment $V(y)$, the expected value of the $(k+1)$th-highest signal conditional on tying with the kth-highest, induces truthful reporting for any given signal s_i, interim. If the first rejected signal is high, winning signals have to be paid subsidies. However, very high subsidy payments are low probability events, as the realizations of all winning signals and the first rejected signal must be high. Further intuition for the surplus-equitable payment is given in the continuation of example 6.12 below. We also note that the surplus-equitable mechanism is ex-post individually rational in the pure private value case. Indeed, $\widetilde{p}_i(s_i, \mathbf{s}_{-i}) = (s_i - G(y)/g(y))\mathbb{1}\{s_i > y\} \leqslant s_i \mathbb{1}\{s_i > y\}$.

We now prove theorem 6.11.

Proof. The payment rule $\widetilde{p}$ results in identical ex-post surpluses of winners. This follows directly from the definition of ex-post surplus under truthful reporting $u_i(s_i, \mathbf{s}_{-i}) = v(s_i, \mathbf{s}_{-i}) - \widetilde{p}_i(s_i, \mathbf{s}_{-i}) = \frac{c}{n}\sum_{j\in[n]} s_j + (1-c)(y - \frac{G(y)}{g(y)}) + V(y)$, for all $s_i > y$, which is independent of the bidder's identity as the first rejected signal is the same for any winner.

Furthermore, the payment $\widetilde{p}$ is interim incentive-compatible. We compute the expected payment

$$\widetilde{P}_i(s_i) = (1-c)\left(s_i G(s_i) - \int_0^{s_i} (yg(y) + G(y))\,\mathrm{d}y\right) + \int_0^{s_i} \widetilde{V}(y,y)g(y)\,\mathrm{d}y = \int_0^{s_i} \widetilde{V}(y,y)g(y)\,\mathrm{d}y.$$

The left-hand term is equal to 0 by integration by parts. Because losers pay nothing, the overall expected payment is equal to the expected payment of winners. The expected utility of a bidder who has signal s_i and reports $\widehat{s}_i$ is given by

$$U_i(s_i, \widehat{s}_i) = \int_0^{\widehat{s}_i} \left(\widetilde{V}(s_i, y) - \widetilde{V}(y,y)\right) g(y)\,\mathrm{d}y, \tag{6.7}$$

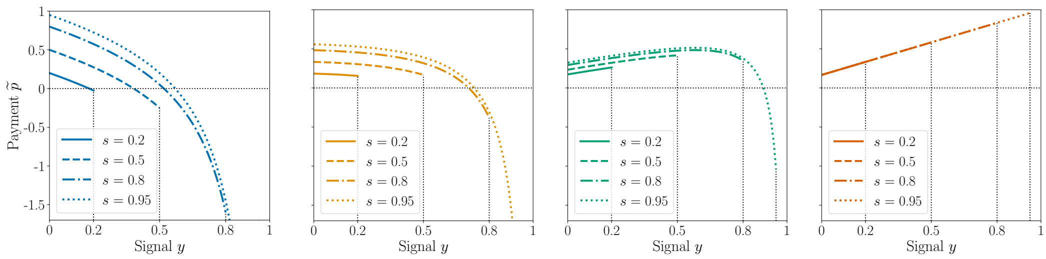
As $\widetilde{V}(s_i, y)$ is increasing in s_i the integrand is positive for $\widehat{s}_i < s_i$ and negative for $\widehat{s}_i > s_i$ and $g > 0$ almost everywhere, the function $U_i(s_i, \widehat{s}_i)$ is uniquely maximized at $\widehat{s}_i = s_i$. □

Distributing surplus equitably ex-post is the strongest of implementations, while an implementation of equal interim surpluses is infeasible in winners-pay auctions. If interim surpluses were equalized across different signals, incentive compatibility cannot hold.[12]

Example 6.12 (continues=uniform-example)**.** We continue the example with $n = 3$ bidders competing for $k = 2$ items and signals distributed uniformly on the support $[0, 1]$. Recall that we have $V(s) = (1 - c)s + \frac{c}{6}(5s + 1)$, and $\frac{G}{g}(y) = \frac{y(2-y)}{2(1-y)}$. Together, we obtain

$$\tilde{p}_i(s, y) = \left((1 - c)\left(s - \frac{y(2 - y)}{2(1 - y)}\right) + \frac{c}{6}(5y + 1)\right)\mathbb{1}\{s > y\}. \tag{6.8}$$

We illustrate the payment for signals $s = 0.2, 0.5, 0.8$ as a function of y in fig. 6.3. Note that each payment corresponding to a signal s is only plotted for $y \leqslant s$, as for $y > s$ the signal s does not win and the payment is zero.



$c = 0$ $\quad$ $c = 0.5$ $\quad$ $c = 0.8$ $\quad$ $c = 1$

Fig. 6.3: Surplus-equitable payments, $\tilde{p}$, for uniform signal distributions as a function of the signal s and the first rejected signal, y, for common value parameters $c \in \{0, 0.5, 0.8, 1\}$.

The payment addresses two countervailing incentives which from the private-value component and the common-value component in the bidders' value structure.

With pure private values, if the first rejected signal is high relative to s, the bidder winning with signal s is paid a subsidy. This subsidy compensates bidders in order to induce them to report truthfully even with high private signals. If a high-signal bidder wins *together* with a low-signal bidder, a large surplus (due to the private signal) is levied in order to equalize ex-post surpluses. Naturally, this would create an incentive to under-report your signal. However, in those cases where all winners' signals are high, no taxation is needed to equalize their surplus, and a subsidy is paid to winning bidders, in order to restore incentive compatibility ex-ante.

[12]The winners-pay assumption is indispensable. Without it, interim surpluses among winners can indeed be equal. In the example with $n = 3$ and $k = 2$, uniform signals and pure private values ($c = 0$), it can be verified that an interim payment of $L(s_i) = s_i(5s_i - s_i^2 - 4)/3(1 - s_i)^2$ charged to the losing bidders achieves equal expected surpluses of the winners of $\frac{2}{3}$. However, then the losing bidders would naturally enter equity considerations.

With a pure common value, the payment is increasing in the first rejected signal y. In this case, no taxation is needed, as the ex-post surpluses are equal under any uniform payment rule. However, truthful reporting must be incentivized, which a second-price rule (adjusted for the common value) achieves. Recall that $V(y)$ is the expected value of a bidder with signal y conditional on tying with the k-th highest among their $n-1$ competitors. With a stronger common value component, the taxation of the surplus due to the private signal is less important. Instead, as seen in fig. 6.3 with $c = 0.8$, the payment is increasing in the first rejected signal. However, for high signals s, there exists an interior maximum of $\widetilde{p}$ in y, a turning point of the countervailing incentives. The taxation of the private signal still hurts the bidder with a higher signal, so a steep subsidy, on a small range of first rejected signals y, must be paid to level incentives in expectation.

The surplus-equitable mechanism relies on the signal prior, which, in practice, is often unknown. However, as we show in the following section, even in the realm of the common uniform and pay-as-bid pricing, it is possible to improve surplus equity. Those results remain, in a large class of signal distributions, prior-free.

6.5 Uniform, pay-as-bid, and mixed auctions

The equity-preferred pricing rule in the mixed auction class crucially depends on the extent of the common value c. As seen for the winners' empirical variance in example 6.10, with uniform signals, for some interior values of c, strictly mixed pricing is optimal. We formalize this fact in section 6.5.1 for any signal distributions. example 6.10 is in line with the general intuition that pay-as-bid pricing may be more equitable with higher private values, and uniform pricing with higher common values. However, as we show in section 6.5.2, this is not true in general. Thus, additional distributional assumptions are necessary for a characterization of equity-preferred auctions. In section 6.5.3, we consider log-concave signal distributions and provide simple and prior-free bounds on the auction that is dominant in pairwise differences.

6.5.1 Equity comparisons in uniform, pay-as-bid, and mixed auctions

We first consider the case of a pure common value ($c = 1$). As every bidder has the same ex-post realized value, ex-post utilities among winners are equalized if everyone pays the same price. This results in pairwise differences in utilities of zero. Once the private value component enters the value function with a non-zero weight, the picture is less clear: it may be pay-as-bid pricing that is dominant in pairwise differences, or it may be some degree of mixed pricing; however, it cannot be uniform pricing. Proofs of this section are relegated to section 6.9.3.

Theorem 6.13. *The uniform-price auction is dominant in pairwise differences iff the common value proportion equals one (pure common value).*

Furthermore, we show that, without any additional assumptions, strictly interior δ-mixed pricing minimizes WEV for a range of common values.

Proposition 6.14. *For any signal distribution, there exists $c^* < 1$, such that for common values in the interval $(c^*, 1)$, there exist δ-mixed auctions with lower* WEV *than pay-as-bid and uniform-price auctions.*

The intuitive notion that uniform pricing equitably distributes surplus under a pure common value may lead us to hypothesize that pay-as-bid auctions are equity-preferred under private values. However, in the following section, we demonstrate a scenario where it fails and show that, with pure private values, uniform pricing can result in lower WEV than pay-as-bid pricing.

6.5.2 Challenging the intuition: private values and uniform pricing

To understand the reversal of the intuition, consider pairwise differences in utility, the building block for WEV. If ex-post absolute differences in utility are greater under uniform pricing than under pay-as-bid pricing for signal pairs with sufficient probability mass, then the reversal may also hold in expectation. To start with, consider any two winning signals $s_i > s_j$, $s_i, s_j \in [0, \bar{s})$ and private values only, i.e., $c = 0$. Let $u_i^0(s_i, \mathbf{s}_{-i})$ and $u_i^1(s_i, \mathbf{s}_{-i})$ denote bidder i's utility in the uniform price and pay-as-bid auction, respectively. Moreover, β^0 and β^1 denote the corresponding symmetric equilibrium bid functions and $Y_{k+1}(\boldsymbol{\beta})$ the first rejected bid. For $\delta \in [0,1]$ and $c = 0$, we have $u_i^\delta(s_i, \mathbf{s}_{-i}) = s_i - \delta\beta^\delta(s_i) - (1-\delta)Y_{k+1}(\boldsymbol{\beta})$. Thus, we have $\Delta u^0 := |u_i^0 - u_j^0| = |s_i - s_j|$ and $\Delta u^1 := |u_i^1 - u_j^1| = |s_i - \beta^1(s_i) - (s_j - \beta^1(s_j))| = |s_i - s_j - (\beta^1(s_i) - \beta^1(s_j))|$. It holds that

$$\Delta U^0 < \Delta U^1 \tag{6.9}$$

$$\Leftrightarrow s_i - s_j < |s_i - s_j - (\beta^1(s_i) - \beta^1(s_j))| \tag{6.10}$$

$$\Rightarrow 2(s_i - s_j) < \beta^1(s_i) - \beta^1(s_j). \tag{6.11}$$

As bid functions are increasing, if $s_i - s_j - (\beta^1(s_i) - \beta^1(s_j))$ was positive, eq. (6.10) could never hold. Thus, eq. (6.11) follows as a necessary condition for uniform pricing to have lower pairwise differences than pay-as-bid pricing. For the same statement to hold for WEV, it must be that the bid function has a slope of at least 2 for a sufficient mass of signals s_i and s_j. Bid function slopes greater than 2 imply that high-signal bidders shade their bids much less, proportionally to their value, than bidders with lower signals. Consequently, the differential in ex-post surplus with pay-as-bid pricing, comparing two sufficiently different signals, are higher than the differential in signals. The latter equals the surplus difference in the uniform-price auction.

The challenge in designing a counter example where WEV is lower in the uniform-price auction than in the pay-as-bid auction is that the slope of β^δ at 0 must be smaller than 1, and thus cannot be greater than 2 for all signals. Thus, the probability mass on regions of the support with a slope of β^δ greater than 2 must be higher, but this also changes β^δ. Nonetheless, the following counter-example satisfies our requirements.

Example 6.15. Consider an auction with n bidders and $k = n - 1$ items. Each bidder i has a pure private value ($c = 0$) given by its signal s_i. The signal is equal to the sum of a Bernoulli random variable with parameter $\varepsilon > 0$ and a random perturbation drawn from Beta$(1, 1/\eta)$, with $\eta > 0$. The resulting signal distribution is continuous, with support $[0, 2]$. This yields the following quantile function:

$$\forall x \in [0,1], \qquad F^{-1}(x) = \mathbb{1}\{x \geqslant \varepsilon\} + \gamma_\eta(x) \qquad \text{where} \qquad \gamma_\eta(x) = \begin{cases} 1 - (1 - \frac{x}{\varepsilon})^\eta & \text{if } x < \varepsilon \\ 1 - \left(1 - \frac{x-\varepsilon}{1-\varepsilon}\right)^\eta & \text{if } x \geqslant \varepsilon. \end{cases}$$

Further derivations and the proof of the below proposition are given in section 6.9.4.

Proposition 6.16. *Let the values be distributed according to the quantile function F^{-1} defined above. For $n \geqslant 5$, there exists η^*, such that for all $\eta \leqslant \eta^*$ it holds that the* winners' empirical variance *under uniform pricing is lower than under pay-as-bid pricing.*

Thus, in order to characterize equity-preferred pricing further, we need additional assumptions. In the next section, we show that, for a large class of signal distributions, simple bounds tell us which auction designs are candidates for being equity-preferred in the class of mixed auctions.

6.5.3 Equity-preferred pricing for log-concave signal distributions

For our subsequent results, we assume a regularity condition on the bidders' signal distributions, *log-concavity*. The family of log-concave distributions contains many common distributions, for example uniform, normal, exponential, logistic or Laplace distributions [BB05].[13]

Definition 6.17. A real-valued function $h \in \mathbb{R}^{\mathbb{R}}$ is *log-concave* if $\log(h)$ is concave.

In this class of signal distributions, simple and prior-free bounds characterize the equity-preferred auction design in the class of mixed auctions. The proofs of theorem 6.18 is developed in section 6.7.

Theorem 6.18. *Let signals be drawn from a log-concave distribution. Then, for a given private value proportion $1-c$, the mixed auction with price discrimination $\delta = 1-c$ is equity-preferred among all mixed auctions with price discrimination of less than $1-c$. Moreover, uniform pricing is dominated in pairwise differences by any strictly mixed pricing with price discrimination of up to $\min\{1, 2(1-c)\}$.*

The equity-preferred pricing rule dominates in pairwise differences all pricing rules with less price discrimination. In other words, theorem 6.18 provides a lower bound on the amount of price discrimination required to rule out dominated mixed auctions. We illustrate theorem 6.18 in fig. 6.4. All pricing rules in the shaded area in red are dominated by the diagonal $1-c$, given any log-concave distribution of bidders' signals.

Moreover, uniform pricing is dominated in pairwise differences by many alternative pricing rules, i.e., these pricing rules are preferred to uniform pricing in terms of equity. This is illustrated in fig. 6.5, in which any pricing rule in the shaded area in green dominates uniform pricing for a given common value c.

The intuition behind theorem 6.18 is simple. As we show in section 6.7, pairwise differences are, for any given common value c, monotonically decreasing in the extent of price discrimination δ as long as δ is between zero and $1-c$. Moreover, we show the equivalence of this result with ex-post utilities that increase in signals. As long as higher signals obtain a higher surplus, more equity can be achieved by taxing higher signals more than lower signal. Because the change in the δ-weighted

[13] Also χ distribution with degrees of freedom $\geqslant 1$, gamma with shape parameter $\geqslant 1$, χ^2 distribution with degree of freedom $\geqslant 2$, beta with both shape parameters $\geqslant 1$, Weibull with shape parameter $\geqslant 1$, and others.

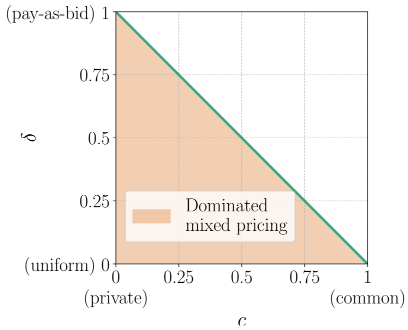


Fig. 6.4: Dominated combinations of c and δ

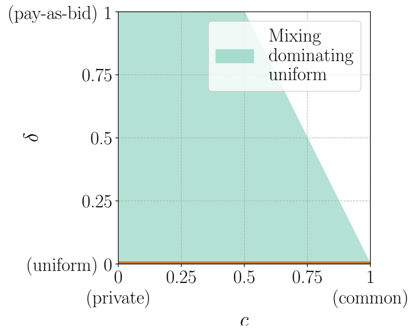


Fig. 6.5: Mixed pricing dominating uniform pricing

bid in δ is increasing in a bidder's signal (as stated in lemma 6.26), increasing the extent of price discrimination will have the desired effect.

A similar intuition explains the dominance of mixed pricing over uniform pricing, where the benefit of higher price discrimination compared to the absence of price discrimination can be realized up to a certain threshold. As long as utilities are increasing in signals, increasing price discrimination results in surplus taxation that benefits equity. We show in section 6.7 that increasing ex-post utilities is equivalent to the slope of equilibrium bid functions being bounded $(1-c)/\delta$. With steeper bid functions, the utilities might decrease in the signals. So, while increasing price discrimination might locally, in a neighborhood of δ, increase pairwise differences, price discrimination is still beneficial compared to uniform pricing. However, for $\delta \geqslant 2(1-c)$, the bid functions are so steep that an increase in price discrimination results in an absolute utility gap between a high signal and a low signal bidder that is greater than under uniform pricing. With such price discrimination, the higher signal bidder is worse off than the low signal bidder.

With theorem 6.18, we can now revisit the question: In terms of equity, should one use pay-as-bid pricing if bidders' values are pure private values? The answer is yes if the signal distributions are log-concave. Moreover, if the common value is small, pay-as-bid pricing is guaranteed to be more equitable than uniform pricing. We state this formally in the corollary below.

Corollary 6.19. *Assume signals are drawn from a log-concave distribution. Then, for pure private values, pay-as-bid pricing is dominant in pairwise differences, and for a common value $c < \frac{1}{2}$, pay-as-bid pricing dominates uniform pricing in pairwise differences.*

The first part of the corollary follows by setting $c = 0$ and the second part follows by setting $\delta = 1$ in theorem 6.18. Our numerical experiments in section 6.6.1 show that, for $c < \frac{1}{2}$, pay-as-bid pricing in fact minimizes WEV for several common distributions. The intuition in the pure private value case carries through under the qualifying assumption of log-concave signals, and it may fail for very concentrated signal distributions. In the latter case, it is important that sufficient probability mass is gathered around higher signals, inducing a bidding equilibrium in which ex-post utilities are decreasing in signals for sufficiently many signal realizations.[14]

For specific signal distributions, we can extend the region where pairwise differences are monotonically decreasing slightly beyond the diagonal $1 - c$. In particular we show that for uniformly distributed signals any pricing dominant in pairwise differences contains a discriminatory proportion of at least $\frac{2n(1-c)}{2n-c(n-2)}$; and for exponentially distributed signals at least $\frac{2n(1-c)}{2n-c(n-(k+1))}$ (see corollary 6.39).[15] Note that both bounds converge to $\frac{1-c}{1-c/2}$ as the number of bidders goes to infinity (and the number of items k is kept constant). We illustrate the bound for signals uniformly distributed on $[0, 1]$ and $n = 3$ bidders and $k = 2$ items in fig. 6.6 below, together with the equity-preferred pricing in terms of WEV. The figure demonstrates that, for high values of c, this bound may be a good heuristic for the optimal pricing rule.

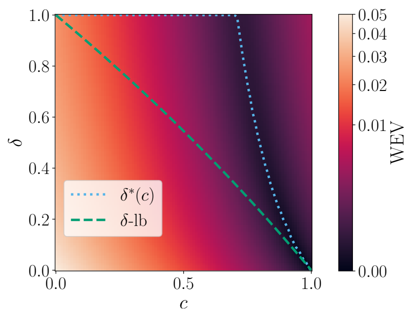


Fig. 6.6: Lower bound δ-lb for candidates dominant in pairwise differences and WEV-optimal pricing δ^*

6.6 Discussion

Our main results hold for all equity metrics that are based on pairwise differences, and, as discussed in the Introduction and in section 6.3, the winners' empirical

[14] For example, with a β-distribution as steep as illustrated in fig. 6.7, section 6.6, clearly violating log-concavity, pay-as-bid pricing is still optimal for a range of common values including pure private values.

[15] The proof should be read in conjunction with theorem 6.18 as it follows a similar reasoning.

variance is particularly attractive as an aggregated one-dimensional metric. We illustrate WEV further in a series of numerical experiments and discuss how it relates to the within-bidder variation of surplus, as well as the empirical variance of surplus between all bidders. We also explain why the regularity assumption of log-concavity is necessary for our argument.

6.6.1 Numerical experiments

We further illustrate the effect of the common value on surplus equity by presenting several numerical examples. Similarly to fig. 6.6, we compute the WEV-minimal pricing $\delta^*(c)$ for any given proportion of the private-common value c. We also illustrate bounds for WEV-minimal pricing and the condition of monotone ex-post utility (MEU). All of our experiments are based on equilibrium bid functions, whose calculation is computationally very expensive. Thus, we rely on theoretical simplifications, such as lemma 6.31 and lemma 6.40 (section 6.9.7). The simulations are performed through numerical integration of our analytical formulae.[16] Finally, some quantities (such as bidding functions) have multiple analytical expressions, among which we choose the most appropriate for accuracy and speed, depending on the value of the signal (e.g., eq. (6.4) can be integrated more efficiently than eq. (6.17), but is less accurate for small signals). Our code is available on github.

We consider a truncated exponential, a truncated normal distribution (both log-concave), and a Beta distribution with shape parameters $(0.5, 0.5)$, which is not log-concave. WEV-minimal pricing, a lower bound on the minimizer, and combinations of common value shares and mixed pricing for which MEU holds are shown in fig. 6.7 for a market with $n = 10$ bidders and $k = 4$ items.

For the truncated exponential distribution, we show the lower bound of $\frac{2n(1-c)}{2n-c(n-(k+1))}$ on $\delta^*(c)$ (cf., corollary 6.39), and for the truncated normal distribution we show the general lower bound $1 - c$ (cf., theorem 6.18). Each of these bounds dominates any extent of price discrimination below it. Note that for the Beta distribution, we cannot provide a theoretical lower bound on the WEV-minimal design δ^*, as the distribution is not log-concave. However, the region where MEU holds can be determined numerically, and its "frontier" provides a lower bound for the WEV-minimal design δ^*. Illustrating this for all three distributions, we observe that the

[16]The efficiency and accuracy of the code rely on various techniques. Most importantly, we rewrite all multidimensional expectations as nested one-dimensional integrals (with variable bounds), which we compute by integrating polynomial interpolations. Second, the code ensures that each quantity is computed at most once, using memorization. Integration is not computationally heavy at all and achieves high precision.

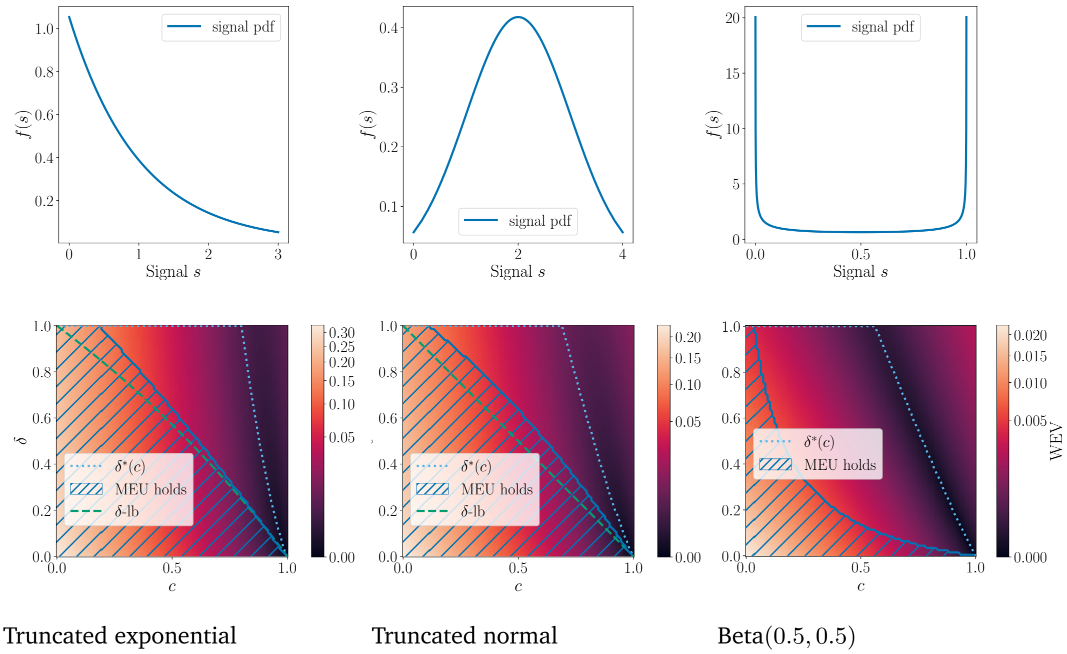


Fig. 6.7: WEV-minimizing design $\delta^*(c)$, monotone ex-post utility (MEU), and lower bounds on $\delta^*(c)$ (δ-lb) for truncated exponential, truncated normal, and Beta$(0.5, 0.5)$ signal distributions

area is much smaller for the Beta distribution. However, MEU is only a sufficient condition for the monotonicity of WEV (while it is necessary and sufficient for the monotonicity of pairwise differences). From the heat maps in fig. 6.7, it is evident that WEV is monotone in δ for any given c up to δ^*.

Finally, we show the WEV-minimal pricing rule $\delta^*(c)$ for each signal distribution. The curve is qualitatively similar in each plot. In line with theorem 6.18 — noting that the exponential and normal distribution are log-concave — the figure illustrates that with a high private value component (low c), pay-as-bid pricing ($\delta = 1$) minimizes WEV; with higher common value components (high c), strictly mixed pricing for some $\delta \in (0, 1)$ minimizes WEV (cf., proposition 6.14); and with a pure common value ($c = 1$), uniform pricing ($\delta = 0$) minimizes WEV (cf., theorem 6.13). Analogous interpretations hold for the Beta distribution, although we cannot give theoretical guarantees.

For small common values, MEU holds for any δ and thus pay-as-bid pricing is dominant in pairwise differences (cf., proposition 6.27). Even for larger common value parameters the WEV-minimal pricing is still pay-as-bid, but eventually strictly mixing ($\delta \in (0, 1)$) is required to minimize WEV. For a pure common value, uniform pricing is WEV-minimal regardless of the signal distribution. Notice also that WEV at the minimal δ^* decreases in c. Naturally, with a higher common value share, bidders' values given different signal realizations as well the corresponding bids move closer together, thus explaining smaller differences in utilities (ex-post and in expectation).

6.6.2 Variance and risk preferences

Surplus equity and distributional concerns are distinct from questions of within-agent variation and associated risk preferences. An appropriate measure to assess the latter is, e.g., the ex-ante variance of bidder surplus. While the two notions are distinct, the measures are linked through the covariance (see also lemma 6.9). In addition, for the pure private value setting, we derive the following result. The proof is deferred to section 6.9.8.

Proposition 6.20. *With pure private values ($c = 0$), the pay-as-bid auction minimizes the ex-ante variance of surplus among all standard auctions with increasing equilibrium bid functions.*

Because of revenue equivalence, note that the previous proposition also implies that $\mathbb{E}[u_i^2]$ is minimal in the pay-as-bid auction among standard auctions. The second moment of surplus links the winners' empirical variance and the empirical variance among all bidders, as shown in lemma 6.34 in the Appendix. As a consequence of lemma 6.34, surplus equity rankings with respect to the winners' empirical variance and the empirical variance among all bidders may not be equivalent. However, applying proposition 6.28 to the pure private value case, we have the following corollary:

Corollary 6.21. *Assuming pure private values ($c = 0$), consider any δ-mixed auction, $\delta \in (0, 1]$, and suppose that the equilibrium bid β^δ satisfies $\frac{\partial \beta^\delta}{\partial s} \leqslant \frac{2}{\delta}$ for all signals $s \in [0, \bar{s})$. Then, the empirical variance (among all bidders) is lower for δ-mixed pricing than for uniform pricing.*

Although this result shows that we can extend equity rankings under pure private values to the empirical variance *among all bidders*, this may not hold in the general case.

6.6.3 Beyond log-concave distributions

A crucial ingredient for theorem 6.18 is that the derivative of the equilibrium bid function is bounded by 1, which holds for log-concave distributions by proposition 6.30. In particular, the density of the first rejected signal must be log-concave. In the following, we provide some insights as to why it is difficult to generalize this result beyond log-concave distributions.

For simplicity, consider the pay-as-bid and the uniform-price auction. Considering log-concave signal distributions, we note that log-concavity is equivalent to (A, G) concavity (a generalization of convexity, see [AVV07]), and $\frac{\partial \beta^\delta}{\partial s} \leqslant 1$ is thus equivalent to (A, G) concavity of $s \mapsto \int_0^s G_k^{n-1}$. One idea to extend our results could then be to consider other generalizations of convexity. Considering proposition 6.28, one might attempt to bound the slope of the bid functions by 2. It holds that $\frac{\partial \beta^\delta}{\partial s} \leqslant 2$ is equivalent to (A, H) concavity of the same function where H is the harmonic mean. But contrary to (A, G) concave functions, there are no simple group closure properties that allow for the (A, H) concavity of f to always imply that of $\int_0^s G_k^{n-1}$. Thus, this route of inquiry does not carry fruits.

We also note that conditions similar to MEU such that uniform pricing yields lower pairwise differences (or WEV) than pay-as-bid pricing are much more difficult to attain. Why? If we follow the same main ideas as in the proof of proposition 6.28, a similar condition using the mean value theorem would be that, for all $s \in (0, \overline{s})$, φ is an expansive mapping, translating into $|(1-c) - \frac{\partial \beta^\delta}{\partial s}| \geqslant 1 - c$. As $\frac{\partial \beta^\delta}{\partial s}$ is strictly positive, we must have $\frac{\partial \beta^\delta}{\partial s} > 1 - c$. For signal distributions with bounded density, g_k^{n-1} is close to zero near $\overline{s}$ (this follows from the definition of order statistics), and therefore $\frac{\partial \beta^\delta}{\partial s}$ is close to zero for a nonzero interval of signals. Thus, $\frac{\partial \beta^\delta}{\partial s} > 1 - c$ cannot hold for all signals on the support, and we cannot rely on similar proof techniques to produce the desired conditions.

6.6.4 Multi-unit demand

A natural question is how robust the equity dominance results for pay-as-bid, uniform, and mixed auctions are with respect to the assumption of unit demand. A simple generalization are flat d_i-unit demands for some $d_{>}1$, $i \in [n]$, meaning that bidders have the same constant marginal value for the first d_i items obtained and 0 after. There are two main challenges to extending our results to this setting, both leading to potential equity-efficiency trade-offs. Firstly, each bidder may win a different number of items, and thus equity could be studied with respect to a bidder's total surplus or average per-unit surplus. The latter allows for some efficiency concerns. The second challenge is behavioral concerning the multiplicity of equilibria in uniform-price auctions: [Nou95] show that even for two bidders and two-unit demand, there is no unique symmetric equilibrium as there exists an inefficient, zero-revenue symmetric equilibrium, as well as an efficient, zero-utility symmetric equilibrium.

[Aus+14] show in their Theorem 1 that a necessary condition for the existence of an ex-post efficient equilibrium in multi-unit auctions is that the total supply of items is an integer multiple of a homogeneous demand, i.e., k/d is an integer ("demand divides supply"). Under this strong assumption, our results extend immediately without ambiguity if we consider a mixed pricing between pay-as-bid and Vickrey pricing. Under Vickrey pricing, a bidder who wins m items must pay the sum of the m highest losing bids, not including the bidder's own losing bids. This payment rule is exactly that of the VCG mechanism; hence, the Vickrey auction is incentive-compatible and efficient. When demand divides supply supply, all winners of the Vickrey auction will receive exactly d items, and each winner pays the same price, since they face the same losing bids. This renders the Vickrey payment into a uniform-pricing rule. Under pay-as-bid pricing, the equilibrium bids for the first d-items are the same as for the 1-unit demand case, and thus each bidder also receives exactly d items (see also [Aus+14]). The consequence of demand dividing supply in both the pay-as-bid and Vickrey auction formats is that items are sold in bundles of size d, with payments and values scaled by d. Therefore, all results of the unit-demand case hold with flat d-unit demand when d divides k.

When demand does not divide supply, our results do not extend to the multi-unit case. In the example in [Aus+14] with two bidders with flat two-unit demand and pure private values, there exists an inefficient equilibrium in the uniform-price auction. In this equilibrium, each bidder is allocated one unit, which is always more surplus-equitable than the efficient equilibrium of the pay-as-bid auction which allocates both items to the higher-value bidder.

6.6.5 Revenue maximization and reserve prices

Standard auctions lead to an efficient allocation of items in our model. Moreover, for a large class of probability distributions (and common values), these auctions are also revenue maximizing (see [BK96]). They show that if, in addition to independent signals, the bidders' marginal revenues are increasing in signals and weakly positive, standard auctions are revenue maximizing. With independent signals, the marginal revenue of bidder i given a realization of signals s is given by $MR_i(s_i) = \frac{-1}{f(s_i)} \frac{\partial}{\partial s_i} (v(s_i)(1 - F(s_i))$. Marginal revenues are positive in many applications where bidders' values are high, i.e., the support of signals is sufficiently positive.[17] In those cases, it is optimal for the seller to sell their entire supply. Our class of auctions is also revenue maximizing if the seller is legally required to sell their entire supply. In practice, either (or both) of the two conditions are observed in many high-stake auctions, e.g., for spectrum licenses.

More generally, seller revenue can be maximized at the expense of efficiency [RS81; Mye81]. A common tool to raise revenue is to set a reserve price $r > 0$ such that only bids exceeding r can win and pay the auction price, or at least r, whichever is higher. [Mye81; RS81] show that, with pure private values, an optimal reserve price maximizes the seller's expected revenue.

Some of our results on equity-preferred pricing extend to pay-as-bid and uniform-price auctions with an additional reserve price. With a reserve price, the number of winners, although identical in different standard auctions with identical reserve prices, is not necessarily equal to the number of items k and becomes a random variable. In the uniform-price auction, the standard derivation leads to an equilibrium bid of $V(s)$ for $s > s_r$, where s_r is some threshold, and 0 otherwise. Using revenue equivalence, we obtain the equilibrium bid in the pay-as-bid auction $\beta_r^{\delta=1}(s) = (r - V(s_r))G(s_r)/G(s) + V(s) - \int_{s_r}^{s} V'(y)G(y)dy/G(s)$ for $s > s_r$ and 0 otherwise. For the equity comparison of pay-as-bid and uniform pricing, we establish the following property.

Proposition 6.22. *If monotone ex-post utility (MEU, definition 6.24) holds in the pay-as-bid auction without a reserve price, then it also holds with a strictly positive reserve price.*

The proof is given in section 6.9.8. Then, the following corollary is immediate (a slightly weaker statement than proposition 6.27).

Corollary 6.23. *Suppose MEU holds in the equilibrium of the pay-as-bid auction without reserve price. Then the pay-as-bid auction with a given reserve price is equity-preferred to (dominates in pairwise differences) any uniform-price auction with the same reserve price.*

[17] In our model, without loss of generality, the signal support includes the lower bound zero.

Conclusion

This article studies the division of surplus between bidders in auction design. We introduce a family of equity measures, based on absolute pairwise differences in realized utilities, that includes popular metrics such as the empirical variance and the expected Gini index. Considering standard and winners-pay auctions in an independent signal setting with single-crossing values, the equity design objective is costless in terms of potential trade-offs with efficiency and revenue.

First, we design the surplus-equitable mechanism, a direct and truthful mechanism that efficiently allocates the items for sale and equalizes the winners' realized utilities. Each winner a personalized price, while losers pay nothing. Turning to the class of uniform, pay-as-bid, and mixed auctions, we show that, in most cases, some degree of price discrimination is beneficial in terms of equity.The equity-preferred auction design crucially depends on the common value proportion in the bidders' value structure. Our results also have substantial implications for the design of multi-unit auctions in practice. By carefully selecting a pricing mixture based on (an estimate of) the common value, auctioneers can achieve a more equitable division of surplus among winning bidders.

Future research could explore the trade-offs between efficiency, revenue, and equity, or extend our analysis to other types of auctions and value distributions. For example, in multi-unit demand settings in which items may be allocated inefficiently (section 6.6.4), trade-offs become relevant. In practice, other designs such as dynamic auctions or the Spanish auction [ÁM07] are used, and understanding the impact of such designs on surplus equity remains an open question.

6.7 Structural insights and proofs of theorem 6.18

In this section, we provide an overview of the proof of theorem 6.18. We combine two propositions on monotonicity of pairwise differences and dominance of pairwise differences, respectively, with a third proposition that bounds the slope of bid functions. In particular, we identify the property of *monotone ex-post utility* as a fundamental and sufficient condition for our dominance results.

Definition 6.24 (Monotone ex-post utility)**.** The ex-post utility $u(\mathbf{s})$ satisfies *monotone ex-post utility (MEU)* iff, for any two signals $s_i, s_j \in [0, \overline{s})$ and $\forall\ \mathbf{s}_{-i}, \mathbf{s}_{-j}$, $s_i \leqslant s_j \Leftrightarrow u_i(s_i, \mathbf{s}_{-i}) \leqslant u_j(s_j, \mathbf{s}_{-j})$.

Monotone ex-post utility (MEU) relates to the slope of equilibrium bids as follows.

Lemma 6.25. *An equilibrium satisfies monotone ex-post utility iff equilibrium bid functions β^δ satisfy $\frac{\partial \beta^\delta}{\partial s} \leqslant \frac{1-c}{\delta}$ for all signals $s \in [0, \overline{s})$.*

Proof: See section 6.9.5.

The ex-post difference in utilities depends only on the private value proportion $(1-c)s$ and the discriminatory part of the payment $\delta\beta^\delta$. Thus, as long as the discriminatory payment does not grow faster in the signal than the private-value share, ex-post utilities are monotone.

As seen in fig. 6.1, the equilibrium exhibit several monotonicity properties, and these hold beyond uniform signals. By assumption, equilibrium bids are increasing in the bidder's own signal. Equilibrium bids are also decreasing in the extent of price discrimination: if a higher proportion of one's own bid affects the price, the incentive to bid-shade increases. Finally, the change in the payment corresponding to a bidder's own bid, due to a signal increase, is increasing in the weight of price discrimination, and vice versa. The latter monotonicity is crucial for proposition 6.27.

Lemma 6.26. *The equilibrium bid functions satisfy the following monotonicity properties:*

1. *$\beta^\delta(s)$ is strictly increasing in s, for all fixed $\delta \in [0,1]$ (consistent with the assumption), and is strictly decreasing in δ, for all fixed $s \in (0, \overline{s})$.*
2. *$\frac{\partial(\delta\beta^\delta(s))}{\partial\delta}$ is strictly increasing in s, for all fixed $\delta \in [0,1]$, and $\frac{\partial(\delta\beta^\delta(s))}{\partial s}$ is strictly increasing in δ, for all fixed $s \in (0, \overline{s})$.*

Proof: See section 6.9.2.

We now characterize the fundamental role of monotone ex-post utility: it is equivalent to the monotonicity property of pairwise differences.

Proposition 6.27. *For a given common value c and for some $\bar{\delta} \in [0,1]$, pairwise differences are monotonically decreasing over $[0, \bar{\delta}]$ if and only if the equilibrium (which depends on c and $\bar{\delta}$) satisfies MEU.*

Proof: See section 6.9.5.

The equivalence between decreasing pairwise differences and MEU being satisfied in equilibrium is crucial in our proof of theorem 6.18. When MEU holds, the slope of the equilibrium bid function is sufficiently flat and more price discrimination impacts higher signal bidders more than lower signal bidders. In contrast, including more uniform pricing in the price mix will, proportionally to the change in δ, offer higher signal bidders a higher discount than lower signal bidders and thus does not improve surplus equity. A similar intuition holds for proposition 6.28 below (for details on the intuition, see section 6.5.3).

Proposition 6.28. *For a given common value c, consider any δ-mixed auction, $\delta \in (0,1]$, and suppose the equilibrium bidding function β^δ satisfies $\frac{\partial \beta^\delta}{\partial s} \leqslant \frac{2(1-c)}{\delta}$ for all signals $s \in [0, \bar{s})$. Then, δ-mixed pricing dominates uniform pricing in pairwise differences.*

Proof: See section 6.9.5.

Example 6.29 (continues=uniform-example)**.** Whether MEU is satisfied can be verified numerically, either by computing differences in realized utilities for every pair of signals or simply by checking the derivative of the bid function. We illustrate this for the example of uniform signal distributions and $n = 3$ and $k = 2$ in fig. 6.8 below. For example, with $c = 0.8$ and $\delta = 0.3$, close to the MEU boundary in fig. 6.8, the derivative of the bid function cannot be larger than $0.667 = \frac{1-0.8}{0.3}$. From fig. 6.1, the slope of the bid function with $c = 0.8$ and $\delta = 0.3$ is close to 0.68 for low signals. Thus, for this combination of c and δ, MEU is not satisfied.

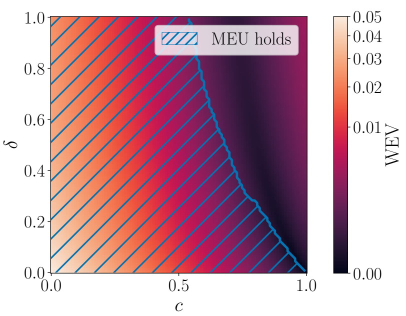


Fig. 6.8: Monotone ex-post utility for common value c and price discrimination δ

The final crucial proposition bounds the slope of the equilibrium bid functions by 1 for the family of log-concave signal distribution.

Proposition 6.30. *If the signal density f is log-concave, then $\frac{\partial \beta^\delta(s)}{\partial s} \leqslant 1$ for all signals $s \in [0, \overline{s})$.*

The proof is given in section 6.7.1 below. With pure private values, this bound implies that ex-post utility is non-decreasing in signals for log-concave signal distributions. Indeed, for $u(s) = s - \delta\beta^\delta(s) - (1-\delta)Y_{k+1}(\boldsymbol{\beta})$, we have that $\frac{\partial u}{\partial s} = 1 - \delta\frac{\partial \beta^\delta(s)}{\partial s} \geqslant 0$. A similar reasoning leads to theorem 6.18.

Proof of theorem 6.18. Because of proposition 6.30, we have that under log-concave signal distributions MEU holds if $\delta \leqslant (1-c)$, as $\frac{\partial \beta^\delta}{\partial s} \leqslant 1 \leqslant \frac{(1-c)}{\delta}$ (see lemma 6.25). Thus, applying proposition 6.27, pairwise differences are monotonically decreasing for $\delta \in [0, 1-c]$ and theorem 6.18 follows.

Similarly, because of proposition 6.30, it holds that with log-concave signals, $\frac{\partial \beta^\delta}{\partial s} \leqslant 1 \leqslant \frac{2(1-c)}{\delta}$ if $\delta \leqslant 2(1-c)$. Applying proposition 6.28, it follows that any mixed pricing with $\delta \in (0, 2(1-c)]$ dominates uniform pricing in pairwise differences. □

6.7.1 Proving the bound on bid function slopes

Bounding the bid function slope for log-concave distributions requires three main observations, which we detail in the lemmas below and then use to prove proposition 6.30.

The first lemma establishes a simplified expression of $V(s)$ which allows to bound $V'(s)$ by 1.

Lemma 6.31. *Assuming a pure common value ($c = 1$), $V(s)$ is differentiable on $(0, \overline{s})$, and can be expressed as*

$$V(s) = \frac{2}{n}s + \frac{n-k-1}{n}\frac{\int_0^s tf(t)\,\mathrm{d}t}{F(s)} + \frac{k-1}{n}\frac{\int_s^{\overline{s}} tf(t)\,\mathrm{d}t}{1-F(s)}$$

Moreover, if the signal density f is log-concave, then $V'(s) \leqslant 1$ for all signals $s \in [0, \overline{s})$.

Proof: See section 6.9.6.

The proof proceeds by noticing that order statistics conditioned on other order statistics behave just like order statistics of a truncation of the original distribution. Thus, a more tractable expression of the expected valuation V can be derived for the pure common value case ($c = 1$). Together with results on log-concavity by [BB05], we use this expression to show that $V' \leqslant 1$ for all signals $s \in (0, \overline{s})$.

The next lemma establishes a sufficient condition for the equilibrium bid function slope to be bounded by 1 in the pure private value case. Differentiating twice $\int_0^s G^{1/\delta}$, we establish that its log-concavity is equivalent to $\frac{\partial \beta^\delta(s)}{\partial s} \leqslant 1$.

Lemma 6.32. *Assuming private values ($c = 0$), for any $\delta \in (0, 1]$, $\frac{\partial \beta^\delta(s)}{\partial s} \leqslant 1$ iff $\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y$ is log-concave.*

Proof: See section 6.9.6.

Finally, we establish that a log-concave signal density is sufficient for the integral of their order statistics to be log-concave, using closure properties of product and integration of log-concave distributions, and results by [BB05].

Lemma 6.33. *If the density of signals f is log-concave, then so is $\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y$.*

Proof: See section 6.9.6.

With the three lemmas above, we can prove proposition 6.30.

Proof of proposition 6.30. First, we recall the expression of the derivative of the bid function for any $s \in (0, \overline{s})$:

$$\frac{\partial \beta^\delta(s)}{\partial s} = \frac{g(s)}{G(s)} \frac{\int_0^s V'(y) G^{\frac{1}{\delta}}(y)\,\mathrm{d}y}{\delta G^{\frac{1}{\delta}}(s)} \tag{6.12}$$

Note that for any $c \in [0, 1]$, $V(s)$ is a linear combination of s and V. In the case of a pure common value, the derivative of the latter is bounded by 1 by lemma 6.31. Hence for any c, $V'(s) \leqslant 1$. Moreover, because of lemma 6.33, we know that $\int_0^s G^{\frac{1}{\delta}}$ is log-concave, and we can therefore apply lemma 6.32. Hence using the above results,

$$\frac{\partial \beta^\delta(s)}{\partial s} \leqslant \frac{g(s)}{G(s)} \frac{\int_0^s \max_t V'(t) G^{\frac{1}{\delta}}(y)\,\mathrm{d}y}{\delta G^{\frac{1}{\delta}}(s)} \leqslant \frac{g(s)}{G(s)} \frac{\int_0^s 1 \cdot G^{\frac{1}{\delta}}(y)\,\mathrm{d}y}{\delta G^{\frac{1}{\delta}}(s)} \leqslant 1. \tag{6.13}$$

□

6.8 Revenue equivalence and efficiency

We recall results from [Kri09] that show that the auctions we consider exhibit revenue equivalence and (allocative) efficiency.

Proposition (Revenue equivalence, [Kri09])**.** *Assuming iid signals, any standard auction, under any symmetric and increasing equilibrium with an expected payment of zero at value zero, yields the same expected revenue to the seller.*

We note that the crucial assumption for revenue equivalence is the independence of signals. In settings where signals are correlated, revenue equivalence fails [Kri09, Chapter 6.5].

We can extend this proposition to our setting: multi-unit single unit demand with common value and i.i.d. signals.

Proposition (Revenue Equivalence)**.** *Assuming i.i.d. signals, common value, single unit demand, any standard and winners pay auction with has equal revenue.*

Proof. The interim utility is given by

$$U_i(s_i, \widehat{s}_i) = \mathbb{E}_{y=Y_k(\mathbf{s}_{-i})}[\mathbb{1}\{\widehat{s}_i \geqslant y\}\,\widetilde{V}(s_i, y)] - P_i(\widehat{s}_i). \tag{6.14}$$

Letting $G := G_k^{n-1}$ and $g := g_k^{n-1}$, we obtain $\partial_1 U_i(s_i, \widehat{s}_i) = \int_0^{\widehat{s}_i} \partial_1 \widetilde{V}(s_i, y) g(y)\,\mathrm{d}y$. Note that the expression is simple because, although values are not private, signals are independent. Let $U_i(s_i) = \max_{\widehat{s}_i} U_i(s_i, \widehat{s}_i)$ in the direct incentive-compatible mechanism, in which the maximum is obtained at $\widehat{s}_i = s_i$ due to incentive compatibility. Then, by the envelope theorem, we must have $U_i'(s_i) = \partial_1 U_i(s_i, s_i)$. Thus, we have $U_i(s_i) = U_i(0) + \int_0^{s_i} U_i'(x)\,\mathrm{d}x$, and consequently $P_i(s_i) = \int_0^{s_i} \widetilde{V}(s_i, y) g(y)\,\mathrm{d}y - U_i(0) - \int_0^{s_i} U_i'(x)\,\mathrm{d}x$. From *winners-pay* and the continuity of the signals follows $U_i(0) = P_i(0) = 0$. ☐

In particular, all the mixed auctions considered in the paper, as well as the equitable direct mechanism, have the same revenue.

A value function $v(\mathbf{s})$ satisfies the *single crossing* condition if for all $i, j \neq i \in [n]$ and for all s, $\frac{\partial v(s_i, \mathbf{s}_{-i})}{\partial s_i} \geqslant \frac{\partial v(s_j, \mathbf{s}_{-j})}{\partial s_i}$, and the value function v as given in assumption 6.1 satisfies this condition.

Proposition (Efficiency, [Kri09])**.** *Any standard auction, under any symmetric and increasing equilibrium and values satisfying the single-crossing condition, is efficient.*

Given the prior propositions, we can focus on the question of surplus distribution among bidders more succinctly without considering potential trade-offs.

6.9 Proofs

6.9.1 Surplus equity

Proof of lemma 6.9. The empirical variance of surplus can be transformed as follows.

$$\begin{aligned}
\mathbb{E}_s\left[\frac{1}{n-1}\sum_{i}^{n}\left(u_i-\frac{1}{n}\sum_{j}^{n}u_j\right)^2\right] &= \mathbb{E}_s\left[\frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(u_i-u_j)^2\right] \\
&= \frac{\mathbb{E}_s\left[(u_1-u_2)^2\right]}{2} \\
&= \mathbb{E}_s[u_1^2]-\mathbb{E}_s[u_1u_2] \\
&= \mathrm{Var}(u_1)-\mathrm{Cov}(u_1,u_2)
\end{aligned}$$

Similarly, the empirical variance conditioned on winning can be written as

$$\begin{aligned}
\mathbb{E}_s\left[\frac{1}{k-1}\sum_{i=1}^{k}\left(u_i-\frac{1}{k}\sum_{j=1}^{k}u_j\right)^2\middle|1,\dots,k\text{ win}\right] &= \frac{\mathbb{E}_s\left[(u_1-u_2)^2\mid 1\text{ and }2\text{ win}\right]}{2} \\
&= \mathbb{E}_s\left[u_1^2\mid 1\text{ wins}\right]-\mathbb{E}_s\left[u_1u_2\mid 1\text{ and }2\text{ win}\right] \\
&= \mathrm{Var}(u_1\mid 1\text{ wins})-\mathrm{Cov}(u_1,u_2\mid 1\text{ and }2\text{ win}).
\end{aligned}$$

$\square$

With pure private values, ex-post individual rationality holds. The lemma below shows that, in this case, any ranking of auction formats in terms of ex-ante variance (Var) or winners' ex-ante variance (WV) is identical. In contrast, a ranking with respect to empirical variance (EV) can differ depending on whether only winners are considered or all bidders.

Lemma 6.34. *Assuming that the auction is a* winners-pay auction*, the empirical variance and the ex-ante variance can be decomposed, respectively, as* EV $= \frac{k(k-1)}{n(n-1)} \cdot$ WEV $+ \left(1 - \frac{k-1}{n-1}\right) E_{\mathbf{s}}[u_1^2]$ *and* Var $= \frac{k}{n} \cdot$ WV $+ \left(\frac{n}{k} - 1\right) \cdot E_{\mathbf{s}}[u_1]^2$.

Recall that $E_{\mathbf{s}}[u_1]$ does not depend on the auction format (by revenue equivalence), while $E_{\mathbf{s}}[u_1^2]$ does.

Proof of lemma 6.34. We first note that

$$\mathrm{WV} = E_{\mathbf{s}}[u_1^2 \,|\, 1 \text{ wins}] - E_{\mathbf{s}}[u_1 \,|\, 1 \text{ wins}]^2 = \frac{n}{k} E_{\mathbf{s}}[u_1^2] - \left(\frac{n}{k}\right)^2 E_{\mathbf{s}}[u_1]^2$$

For the ex-ante variance, we write:

$$\begin{aligned}
\mathrm{Var} &= E_{\mathbf{s}}[u_1^2] - E_{\mathbf{s}}[u_1]^2 = P_{\mathbf{s}}[1 \text{ wins}] \cdot E_{\mathbf{s}}[u_1^2 \,|\, 1 \text{ wins}] - E_{\mathbf{s}}[u_1]^2 \\
&= P_{\mathbf{s}}[1 \text{ wins}] \cdot E_{\mathbf{s}}[u_1^2 \,|\, 1 \text{ wins}] - P_{\mathbf{s}}[1 \text{ wins}] \cdot E_{\mathbf{s}}[u_1 \,|\, 1 \text{ wins}]^2 + P_{\mathbf{s}}[1 \text{ wins}] \cdot E_{\mathbf{s}}[u_1 \,|\, 1 \text{ wins}]^2 - E_{\mathbf{s}}[u_1]^2 \\
&= P_{\mathbf{s}}[1 \text{ wins}] \cdot \mathrm{WV} + P_{\mathbf{s}}[1 \text{ wins}] \cdot E_{\mathbf{s}}[u_1 \,|\, 1 \text{ wins}]^2 - E_{\mathbf{s}}[u_1]^2 \\
&= P_{\mathbf{s}}[1 \text{ wins}] \cdot \mathrm{WV} + \frac{P_{\mathbf{s}}[1 \text{ wins}]^2}{P_{\mathbf{s}}[1 \text{ wins}]} \cdot E_{\mathbf{s}}[u_1 \,|\, 1 \text{ wins}]^2 - E_{\mathbf{s}}[u_1]^2 \\
&= P_{\mathbf{s}}[1 \text{ wins}] \cdot \mathrm{WV} + \left(\frac{n}{k} - 1\right) \cdot E_{\mathbf{s}}[u_1]^2
\end{aligned}$$

For the empirical variance, we write:

$$\begin{aligned}
\mathrm{WEV} &= E_{\mathbf{s}}[u_1^2] - E_{\mathbf{s}}[u_1 u_2] = P_{\mathbf{s}}[1 \text{ wins}] \cdot E_{\mathbf{s}}[u_1^2 \,|\, 1 \text{ wins}] - P_{\mathbf{s}}[1 \text{ and } 2 \text{ win}] \cdot E_{\mathbf{s}}[u_1 u_2 \,|\, 1 \text{ and } 2 \text{ win}] \\
&= P_{\mathbf{s}}[1 \text{ and } 2 \text{ win}] \cdot \mathrm{WEV} + \left(1 - \frac{P_{\mathbf{s}}[1 \text{ and } 2 \text{ win}]}{P_{\mathbf{s}}[1 \text{ wins}]}\right) \cdot E_{\mathbf{s}}[u_1^2]
\end{aligned}$$

□

Proof of proposition 6.8. Without loss of generality, consider an outcome profile u with three outcomes, u_i, u_j and U, where $u_i > u_j$, and U is arbitrary. Induce a Pigou-Dalton transfer $t > 0$ such that $u_i' = u_i - t > u_j$ and $u_j' = u_j + t < u_i$, and U remains the same. The outcome profile after the transfer is denoted u'. We show that the ranking between u and u' according to WEV coincides with what the Pigou-Dalton principle requires, namely it must be that $WEV(u') <$ WEV(u). Let $W := (u_i - U)^2 + (u_j - U)^2$. Then

$$\begin{aligned}
&(u_i' - U)^2 + (u_j' - U)^2 \\
= &(u_i - t - U)^2 + (u_j + t - U)^2
\end{aligned}$$

$$
\begin{aligned}
&= (u_i - U)^2 - 2t(u_i - U) + t^2 + (u_j - U)^2 + 2t(u_j - U) + t^2 \\
&= W + 2t\,(t - u_i + U + u_j - U) \\
&= W + 2t\,(u_j - (u_i - t)) \\
&< W
\end{aligned}
$$

The final inequality follows by the assumption that the transfer does not make i poorer than j was to start with. As U was arbitrarily chosen and, to compute WEV, expectations are taken around the sum of squared differences of the realized utilities, the result follows. □

6.9.2 Equilibrium bidding

Proof of proposition 6.3. Consider bidder i and let all bidders $j \neq i$ follow the bidding strategy $\beta^U(s_j) = \widetilde{V}(s_j, s_j)$. First, observe that β^U is continuous and increasing. Then bidder i's expected payoff when their signal is s_i and bidding $\beta^U(z)$ is given by

$$
U(s_i, z) := \int_0^z \left(\widetilde{V}(s_i, y) - \widetilde{V}(y, y)\right) g_k^{n-1}(y)\,\mathrm{d}y
$$

Because $\widetilde{V}(s_i, y)$ is increasing in s_i, it holds for all $y < s_i$ that $\widetilde{V}(s_i, y) - \widetilde{V}(y, y) > 0$, and for all $y > s_i$ that $\widetilde{V}(s_i, y) - \widetilde{V}(y, y) < 0$. Therefore, choosing $z = s_i$ maximizes bidder i' expected payoff $U(s_i, z)$. □

Proof of proposition 6.4. First, observe that β^δ is continuous. We verify that it is also monotone: writing $G_k^{n-1} =: G$, $g_k^{n-1} =: g$, and $\widetilde{V}(s, s) =: V(s)$, an alternative expression for β^δ is given by

$$
\beta^\delta(s) = V(s) - \frac{\int_0^s V'(y) G(y)^{\frac{1}{\delta}}\,\mathrm{d}y}{G(s)^{\frac{1}{\delta}}}. \tag{6.15}
$$

In particular, it is differentiable almost everywhere and we can compute its derivative.

$$
\beta^{\delta\prime}(s) = \frac{g(s) \int_0^s V'(y) G(y)^{\frac{1}{\delta}}\,\mathrm{d}y}{\delta G(s)^{1+\frac{1}{\delta}}} \tag{6.16}
$$

which it positive almost everywhere. Next, assume that all bidders $j \neq i$ follow the bidding strategy β^δ, and let $\beta^\delta(z)$ be bidder i's bid, whose expected utility is given by

$$U(s_i, z) := \int_0^z \left(\widetilde{V}(s_i, y) - \delta\beta^\delta(z) - (1-\delta)\beta^\delta(y)\right) g(y)\,\mathrm{d}y$$

The derivative of $U(s_i, z)$ is

$$\begin{aligned}\frac{\mathrm{d}U}{\mathrm{d}z}(s_i, z) &= \widetilde{V}(s_i, z)g(z) - \delta\beta^{\delta\prime}(z)G(z) - \delta\beta^\delta(z)g(z) - (1-\delta)\beta^\delta(z)g(z)\\ &= (\widetilde{V}(s_i, z) - \beta^\delta(z))g(z) - \delta\beta^{\delta\prime}(z)G(z).\end{aligned}$$

In equilibrium, the first order condition requires $\frac{\mathrm{d}U}{\mathrm{d}z}(s_i, s_i) = 0$. We solve this differential equation using $G^{\frac{1}{\delta}-1}$ as the integrating factor. We obtain

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[G(z)^{\frac{1}{\delta}}\beta^\delta(z)\right] = \left(\frac{1}{\delta}G(z)^{\frac{1}{\delta}-1}\right)\cdot(\beta^\delta(z)g(z) + \delta\beta^{\delta\prime}G(z)) = \left(\frac{1}{\delta}G(z)^{\frac{1}{\delta}-1}\right)\cdot\widetilde{V}(s_i, z)g(z).$$

Solving for β^δ, we obtain

$$\beta^\delta(s) = \frac{\int_0^s V(y)g_k^{n-1}(y)G_k^{n-1}(y)^{\frac{1}{\delta}-1}\,\mathrm{d}y}{\delta G_k^{n-1}(s)^{\frac{1}{\delta}}}. \tag{6.17}$$

Using equations (6.15) and (6.16), and the fact that $\widetilde{V}(s_i, z)$ is increasing in s_i, we have that $\frac{\mathrm{d}U}{\mathrm{d}z}$ is positive when $z \leqslant s_i$ and negative when $z \geqslant s_i$. Therefore, choosing $z = s_i$ maximizes i's expected payoff $U(s_i, z)$.

Finally, we derive the expression for β^δ stated in the proposition from eq. (6.17). Writing $G_k^{n-1} =: G$ and $g_k^{n-1} =: g$, observe that the derivative of $\delta G^{\frac{1}{\delta}}$ is $gG^{\frac{1}{\delta}-1}$. Using integration by parts and a change of variable, we obtain

$$\begin{aligned}\int_0^s V(y)g(y)G(y)^{\frac{1}{\delta}-1}\,\mathrm{d}y &= \left[\delta V(y)G(y)^{\frac{1}{\delta}}\right]_0^s - \delta\int_0^s V'(y)G(y)^{\frac{1}{\delta}}\,\mathrm{d}y\\ &= \delta V(s)G(s)^{\frac{1}{\delta}} - \delta\int_{V(0)}^{V(s)} G(V^{-1}(y))^{\frac{1}{\delta}}\,\mathrm{d}y.\end{aligned}$$

Dividing by $\delta G^{\frac{1}{\delta}}$ gives the result. □

Lemma 6.35. *For any continuous function $\varphi : [0, \overline{s}) \to \mathbb{R}$, and for all $s \in (0, \overline{s})$, we have*

$$\lim_{\delta\to 0}\int_0^s \frac{\varphi(t)}{\delta}\left(\frac{G(t)}{G(s)}\right)^{\frac{1}{\delta}}\mathrm{d}t = \varphi(s)\cdot\frac{G(s)}{g(s)}$$

$$\lim_{\delta\to 0}\int_0^s \log\left(\frac{G(s)}{G(t)}\right)\frac{\varphi(t)}{\delta^2}\left(\frac{G(t)}{G(s)}\right)^{\frac{1}{\delta}} \mathrm{d}t = \varphi(s)\cdot\frac{G(s)}{g(s)}$$

where $G_k^{n-1} =: G$ and $g_k^{n-1} =: g$.

Proof. Fix $\delta > 0$, and let $\psi : (0,1] \to \mathbb{R}$ be a continuous function, such that $\psi(u) = O(1/u)$ when $u \to 0$. Using the change of variable $u = v^\delta$, we have that

$$\int_0^1 \frac{\psi(u)}{\delta} u^{\frac{1}{\delta}}\,\mathrm{d}u = \int_0^1 \psi(v^\delta)v^\delta dv$$
$$\int_0^1 \log(1/u)\frac{\psi(u)}{\delta^2} u^{\frac{1}{\delta}}\,\mathrm{d}u = \int_0^1 \log(1/v)\psi(v^\delta)v^\delta dv.$$

Observe that for all fixed $v \in (0,1]$, and taking $\delta \to 0$, the first (resp. second) integrand converges towards $\psi(1)$ (resp., $\psi(1)\log(1/v)$). We define the constant $M = \sup_{u\in(0,1]} u\psi(u)$, we bound the first integrand by M (resp.the second integrand by $M\log(1/v)$), and we use the theorem of dominated convergence, which gives

$$\lim_{\delta\to 0}\int_0^1 \frac{\psi(u)}{\delta} u^{\frac{1}{\delta}}\,\mathrm{d}u = \int_0^1 \psi(1)\,\mathrm{d}v = \psi(1)$$
$$\lim_{\delta\to 0}\int_0^1 \log(1/u)\frac{\psi(u)}{\delta^2} u^{\frac{1}{\delta}}\,\mathrm{d}u = \int_0^1 \psi(1)\log(1/v)\,\mathrm{d}v = \psi(1)$$

To prove the lemma, observe that with the change of variable $u = \frac{G(t)}{G(s)}$, we have

$$\int_0^s \frac{\varphi(t)}{\delta}\left(\frac{G(t)}{G(s)}\right)^{\frac{1}{\delta}}\mathrm{d}t = \int_0^1 \frac{\psi(u)}{\delta} u^{\frac{1}{\delta}}\,\mathrm{d}u$$
$$\int_0^s \log\left(\frac{G(s)}{G(t)}\right)\frac{\varphi(t)}{\delta^2}\left(\frac{G(t)}{G(s)}\right)^{\frac{1}{\delta}}\mathrm{d}t = \int_0^1 \log(1/u)\frac{\psi(u)}{\delta^2} u^{\frac{1}{\delta}}\,\mathrm{d}u$$

where we define

$$\psi(u) := G(s)\cdot\frac{\varphi(G^{-1}(uG(s)))}{g(G^{-1}(uG(s)))}.$$

Finally, it remains to prove that $\psi(u) = O(1/u)$ when $u \to 0$. First, observe that φ is bounded on $[0,s]$. Second, observe that we have

$$\frac{u}{g(G^{-1}(uG(s)))} = \frac{1}{G(s)}\frac{G(x)}{g(x)},$$

where $x = G^{-1}(uG(s)) \to 0$. Because g is positive and integrable in 0, we have that G/g is bounded. Therefore, the overall limit when $\delta \to 0$ is equal to $\psi(1)$, which concludes the proof. □

Lemma 6.36. *The following formulas can be derived:*

$$\beta^\delta(s) = \begin{cases} V(s) & \text{if } \delta = 0 \\ V(s) - \int_0^s V'(y) \left(\frac{G(y)}{G(s)}\right)^{\frac{1}{\delta}} \mathrm{d}y & \text{if } \delta > 0 \end{cases}$$

$$\frac{\partial(\beta^\delta(s))}{\partial s} = \begin{cases} V'(s) & \text{if } \delta = 0, s > 0 \\ \frac{g(s)}{G(s)} \int_0^s \frac{V'(y)}{\delta} \left(\frac{G(y)}{G(s)}\right)^{\frac{1}{\delta}} \mathrm{d}y & \text{if } \delta > 0 \end{cases}$$

$$\frac{\delta\partial(\beta^\delta(s))}{\partial\delta} = \begin{cases} 0 & \text{if } \delta = 0 \text{ or } s = 0 \\ \int_0^s V'(y) \log\left(\left(\frac{G(y)}{G(s)}\right)^{1/\delta}\right) \left(\frac{G(y)}{G(s)}\right)^{\frac{1}{\delta}} \mathrm{d}y & \text{for } \delta, s > 0 \end{cases}$$

$$\frac{\partial^2(\delta\beta^\delta(s))}{\partial s\partial\delta} = \frac{-g(s)}{\delta G(s)} \int_0^s V'(y) \log\left(\left(\frac{G(y)}{G(s)}\right)^{1/\delta}\right) \left(\frac{G(y)}{G(s)}\right)^{1/\delta} \mathrm{d}y \text{ for } \delta, s > 0$$

Proof. In order to derive the value of these functions at points where they are not directly, defined, we will use the dominated convergence theorem.

(1) Let $s \in (0, \overline{s})$. We first look at $\beta^\delta(s) = V(s) - \int_0^s V'(y) \left(\frac{G(y)}{G(s)}\right)^{\frac{1}{\delta}} \mathrm{d}y$. Let $h(\delta, y)$ be the function under the integral. Because G is increasing, for $y < s$ we have that $G(y)/G(s) < 1$. Hence h is dominated by V', and $\lim_{\delta\to 0} h(\delta, y) = 0$, hence by dominated convergence $\beta^\delta(s) = V(s)$ when $\delta = 0$, and the function is separately continuous over $[0, 1] \times [0, \overline{s})$.

(2) We now consider the derivative of β^δ with respect to s. Let $s \in (0, \overline{s})$. There exists $M > m > 0$ such that $s \in [m, M]$. We focus on the derivative of the integral part:

$$-\frac{\partial}{\partial s} V'(y) \left(\frac{G(y)}{G(s)}\right)^{1/\delta} = V'(y) \frac{g(s) G^{1/\delta}(y)}{G^{1/\delta+1}(s)} \leqslant V'(y) \frac{g(s)}{G(s)} \leqslant V'(y) \frac{\sup_{t\in[m,M]} g(t)}{G(m)},$$

where the $\sup_{t\in[m,M]} g(t)$ is finite as g is continuous. Because V' is integrable, we can use dominated convergence. Using the Leibniz integral rule yields the result. The limit as δ goes to 0 can be computed by applying lemma 6.35.

(3) Let us now compute the derivative of β^δ with respect to δ. Again, we use a dominated convergence property to show that the integral and derivative can be inverted. It is easier to show that this can be done for the function $\delta\beta^\delta(s)$, and we have

$$\frac{\partial\delta\beta^\delta(s)}{\partial\delta} = \beta^\delta + \delta\frac{\partial\beta^\delta(s)}{\partial\delta}.$$

Computing the derivative of $\delta\beta^\delta(s)$, we can recover that of $\beta^\delta(s)$.

Let $h(\delta, y, s) = \delta V'(y)(G(y)/G(s))^{1/\delta}$ be the function under the integral part of $\delta\beta^\delta$. We have

$$\begin{aligned}\frac{\partial h(\delta, y, s)}{\partial \delta} &= V'(y)\left(\frac{G(y)}{G(s)}\right)^{1/\delta} - V'(y)\frac{\delta}{\delta^2}\log\left(\frac{G(y)}{G(s)}\right)\left(\frac{G(y)}{G(s)}\right)^{1/\delta} \\ &= V'(y)\left(\frac{G(y)}{G(s)}\right)^{1/\delta} - V'(y)\log\left(\left(\frac{G(y)}{G(s)}\right)^{1/\delta}\right)\left(\frac{G(y)}{G(s)}\right)^{1/\delta}.\end{aligned}$$

The first part is again dominated by V', which is integrable. Focusing on the second part, we define for $0 < u < w < 1$ the function $\psi(u, w) = (u/w)\log(w/u)$. Note that $0 < s < y < \bar{s}$ implies that for $u = G^{1/\delta}(y)$ and $w = G^{1/\delta}(s)$, we have $0 < u < w < 1$ as G is increasing and takes values in $(0, 1)$ over $(0, \bar{s})$ by definition. Fix w, and take the derivative with respect to u: we obtain that $\psi'(u, w) = (\log(w/u) - 1)/w$ which is positive as long as $u \leqslant w/e$ and negative otherwise. The maximum of ψ for $u < w$ is at $u = w/e$ and $\psi(w/e, w) = 1/e$. This shows the right-hand side of h is smaller than $V'(y)/e$, which is also integrable. Overall by dominated convergence we can invert derivative and integral: $\frac{\partial}{\partial\delta}\int h = \int \frac{\partial}{\partial\delta}h$. Thus

$$\frac{\partial \delta\beta^\delta}{\partial \delta} = V(s) - \int_0^s V'(y)\left(\frac{G(y)}{G(s)}\right)^{1/\delta} dy + \int_0^s V'(y)\log\left(\left(\frac{G(y)}{G(s)}\right)^{1/\delta}\right)\left(\frac{G(y)}{G(s)}\right)^{1/\delta},$$

and we recover

$$\delta\frac{\partial \beta^\delta(s)}{\partial \delta} = \int_0^s V'(y)\log\left(\left(\frac{G(y)}{G(s)}\right)^{1/\delta}\right)\left(\frac{G(y)}{G(s)}\right)^{1/\delta}.$$

Using the same upper bound on ψ, we can show that the integrand of $\delta\frac{\partial\beta^\delta(s)}{\partial\delta}$ is smaller than $V'(y)/e$ which allows for domination both in small δ and small s. By dominated convergence, we obtain that the limit of $\delta\frac{\partial\beta^\delta(s)}{\partial\delta}$ as either δ or s go to 0 is 0.

(4) Finally, let us compute the cross derivative. The integrand of $\frac{\partial\beta^\delta(s)}{\partial s}$ is $h(\delta, y, s) = V'(y)(G(y)/G(s))^{1/\delta}$, and its derivative with respect to delta is $-\frac{1}{\delta}V'(y)\log(G^{1/\delta}(y))G^{1/\delta}(y)$. Because this function is continuous on the open set $(0, 1) \times [0, \bar{s})$, we can apply dominated convergence to show that the order of derivative and integral can be reversed. Therefore

$$\frac{\partial^2 \delta\beta^\delta}{\partial\delta\partial s}$$

$$= \frac{g(s)}{G(s)} \frac{-G^{1/\delta}(s)\frac{1}{\delta}\int_0^s V'(y)\log(G^{1/\delta}(y))G^{1/\delta}(y)\,\mathrm{d}y + \frac{1}{\delta}\log(G^{1/\delta}(s))G^{1/\delta}(s)\int_0^s V'(y)G^{1/\delta}(y)\,\mathrm{d}y}{G^{2/\delta}}$$

$$= \frac{-g(s)}{\delta G(s)} \int_0^s V'(y) \log\left(\left(\frac{G(y)}{G(s)}\right)^{1/\delta}\right)\left(\frac{G(y)}{G(s)}\right)^{1/\delta} \mathrm{d}y.$$

□

Lemma 6.37. *Consider a function $\varphi : [0,1] \times (0,\overline{s}) \to \mathbb{R}_+$, such that*

- *$\varphi_\delta : s \mapsto \varphi(\delta, s)$ is continuous over $(0,\overline{s})$ for all fixed $\delta \in [0,1]$,*
- *$\varphi_s : \delta \mapsto \varphi(\delta, s)$ is continuous over $[0,1]$ for all fixed $s \in (0,\overline{s})$,*
- *either all φ_s's are monotone or all φ_δ's are monotone,*

then φ is jointly continuous in δ and s.

Proof. The proof on the open set $(0,1) \times (0,\overline{s})$ is written in [KD69], and directly generalizes to $\delta = 0$ and $\delta = 1$ given that φ is separately continuous in those points. □

Proof of lemma 6.26. Monotonicity follows from the derivatives computed in Lemma 6.36.

□

6.9.3 Equity comparisons in uniform, pay-as-bid and mixed auctions

Proof of theorem 6.13. To prove the "if" direction, note that for $c = 1$, the realized value is identical for all bidders $i \in [n]$ as $v(\mathbf{s}) = \frac{1}{n}\sum_{j\in[n]} s_j$. Thus, with a uniform price that is identical between bidders, they all have identical surplus. For any $\delta > 0$, the payment differs between the winners at least for some signal realizations.

To prove the "only if" let $\varphi^\delta(s) = (1-c)\cdot s - \delta\beta^\delta(s)$. We then have $(u(s_i) - u(s_j))^2 = (\varphi^\delta(s_i) - \varphi^\delta(s_j))^2$ for two winning bids s_i, s_j (see the proof of proposition 6.28 for details). Thus, it holds that

$$\frac{\partial}{\partial\delta}(\varphi^\delta(s_i)-\varphi^\delta(s_j))^2 = -2(\varphi^\delta(s_i)-\varphi^\delta(s_j))\left(\beta^\delta(s_i) - \beta^\delta(s_j) + \delta\frac{\partial\beta^\delta(s_i)}{\partial\delta} - \delta\frac{\partial\beta^\delta(s_j)}{\partial\delta}\right).$$

Using lemma 6.36, we take the limit of β^δ, $\delta\frac{\partial\beta^\delta(s)}{\partial\delta}$, and $\varphi^\delta(s)$, as δ goes to 0. We have that $(\varphi^\delta(s_i) - \varphi^\delta(s_j)) \to (1-c)(s_i - s_j)$ and $(\beta^\delta(s_i) - \beta^\delta(s_j) + \delta\frac{\partial\beta^\delta(s_i)}{\partial\delta} - \delta\frac{\partial\beta^\delta(s_j)}{\partial\delta}) \to$

$(V(s_i) - V(s_j))$. As V is increasing, the product $(V(s_j) - V(s_i))(s_i - s_j)$ is strictly negative almost surely, which concludes the proof. □

Proof of proposition 6.14. We show that for any $c \in (c^*, 1)$, pay-as-bid pricing does not minimize WEV. From this and the "only if" statement in the proof of theorem 6.13, the result follows. Note that WEV is continuous in c and at $c = 1$ it is strictly lower for uniform pricing ($\delta = 0$) than for pay-as-bid pricing ($\delta = 1$) by theorem 6.13. Thus, by the mean value theorem, there exists an open interval $C = (c^*, 1)$, $c^* < 1$, such that, for any $c \in C$, WEV remains strictly lower under uniform pricing than pay-as-bid pricing. □

6.9.4 Challenging the intuition: private values and uniform pricing

We have to show that it is indeed possible to construct an *equilibrium* bid function with a slope greater than 2 for a sufficient mass of signals. For this, we require an extreme signal distribution where, broadly speaking, signals are equal to zero with probability ε and equal to one with probability $1 - \varepsilon$. However, to compute a Bayes-Nash equilibrium, we need a continuous signal distribution (with respect to the Lebesgue measure, without mass points) with connected support to solve the first-order condition. Thus, we add a small perturbation.

For example 6.15, we consider the order statistics of quantiles $F^{-1}(x)$ and not of signals s. For convenience, we define the following distribution functions and densities.

$$\begin{aligned}
\widetilde{G}(x) &:= G_k^{n-1}(F^{-1}(x)) = 1 - (1-x)^{n-1} & \qquad \widetilde{g}(x) &:= (n-1)(1-x)^{n-2} \\
\widetilde{H}(x) &:= G_{k-1}^{n-2}(F^{-1}(x)) = 1 - (1-x)^{n-2} & \qquad \widetilde{h}(x) &:= (n-2)(1-x)^{n-3}
\end{aligned}$$

We choose a continuous distribution of signals, with support $[0, 2]$, where each signal is given by the sum of a Bernoulli(ε) random variable and a random perturbation drawn from Beta$(1, 1/\eta)$, with $\varepsilon = 0.1/n$ and η a small constant. First, we compute the distribution function F and quantile function F^{-1} of the signal distribution. Using the law of total probabilities, we have

$$\begin{aligned}
\forall s \in [0,2], \qquad F(s) &= P[\text{Bernoulli}(\varepsilon) + \text{Beta}(1, 1/\eta) \leqslant s] \\
&= P[\text{Bernoulli}(\varepsilon) = 0] \cdot P[\text{Beta}(1, 1/\eta) \leqslant s] \\
&\quad + P[\text{Bernoulli}(\varepsilon) = 1] \cdot P[\text{Beta}(1, 1/\eta) \leqslant s - 1].
\end{aligned}$$

Simplifying this expression depending on the value of s, we get

$$\forall s \in [0,2], \qquad F(s) = \begin{cases} \varepsilon \cdot (1-(1-s)^{1/\eta}) & \text{if } s \leqslant 1, \\ \varepsilon + (1-\varepsilon)\cdot(1-(2-s)^{1/\eta}) & \text{if } s \geqslant 1. \end{cases}$$

Then, computing piece-by-piece the inverse of F, we obtain

$$\forall x \in [0,1], \qquad F^{-1}(x) = \begin{cases} 1-(1-\frac{x}{\varepsilon})^{\eta} & \text{if } x \leqslant \varepsilon, \\ 2-\left(1-\frac{x-\varepsilon}{1-\varepsilon}\right)^{\eta} & \text{if } x \geqslant \varepsilon. \end{cases}$$

See Figure 6.9 for the CDF and quantile function of the signals.

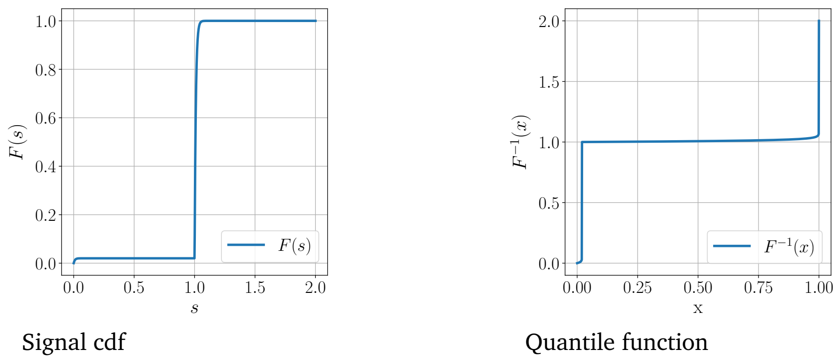


Fig. 6.9: Bidder signals and quantiles for $n = 5$ and $\eta = 0.01$

A bidder with quantile $x \in [0,1]$ bids (truthfully) their signal $F^{-1}(x)$ in the uniform-price auction ($\delta = 0$), which we write as $b_\eta^0(x) := F^{-1}(x) = \mathbb{1}\{x \geqslant \varepsilon\} + \gamma_\eta(x)$, where

$$\forall x \in [0,1], \qquad \gamma_\eta(x) := \begin{cases} 1-(1-\frac{x}{\varepsilon})^{\eta} & \text{if } x < \varepsilon, \\ 1-\left(1-\frac{x-\varepsilon}{1-\varepsilon}\right)^{\eta} & \text{if } x \geqslant \varepsilon. \end{cases}$$

For mixed auctions with $\delta > 0$, the equilibrium bid function is given by Proposition 6.4. Letting $b_\eta^\delta(x) := \beta^\delta(F^{-1}(x))$ denote the equilibrium bid of a bidder with quantile $x \in [0,1]$, we have

$$b_\eta^\delta(x) = \frac{\int_0^{F^{-1}(x)} V(s) g_k^{n-1}(s) G_k^{n-1}(s)^{\frac{1}{\delta}-1}\,\mathrm{d}s}{\delta G_k^{n-1}(F^{-1}(x))} = \frac{\int_0^x F^{-1}(y)\widetilde{g}(y)\widetilde{G}(y)^{\frac{1}{\delta}-1}\,\mathrm{d}y}{\delta \widetilde{G}(x)},$$

where we used the change of variable $y = F(s)$. Finally, using the additive form of F^{-1} we write the equilibrium bid function as $b_\eta^\delta(x) = b_0^\delta(x) + \xi_\eta^\delta(x)$, where

$$\forall x \in [0,1], \qquad b_0^\delta(x) := \frac{\int_\varepsilon^x \widetilde{g}(y)\widetilde{G}(y)^{\frac{1}{\delta}-1}\,\mathrm{d}y}{\delta \widetilde{G}(F^{-1}(x))} = \begin{cases} 0 & \text{if } x < \varepsilon \\ 1 - \left(\frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(x)}\right)^{\frac{1}{\delta}} & \text{if } x \geqslant \varepsilon \end{cases}$$

$$\xi_\eta^\delta(x) := \frac{\int_0^x \gamma_\eta(y)\widetilde{g}(y)\widetilde{G}(y)^{\frac{1}{\delta}-1}\,\mathrm{d}y}{\delta \widetilde{G}(x)}$$

See Figure 6.10 for the bid functions.

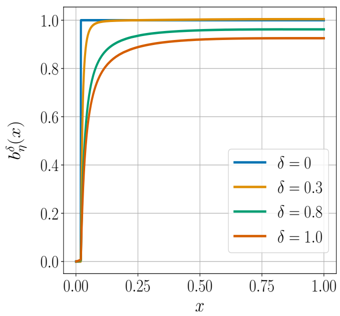


Fig. 6.10: Equilibrium bid as a function of quantiles for $n = 5$ and $\eta = 0.01$

Next, we define the function $\varphi_\eta^\delta(x) := F^{-1}(x) - \delta b_\eta^\delta(x)$, the utility of a winning bidder as a function of their quantile. Denoting WEV_η^δ the *winners' empirical variance* in a δ-mixed auction with noise level η, we write

$$\forall \delta \in [0,1],\ \forall \eta > 0, \qquad \mathrm{WEV}_\eta^\delta = \mathbb{E}_{\mathbf{x}}\left[\frac{(\varphi_\eta^\delta(x_1) - \varphi_\eta^\delta(x_2))^2}{2} \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right].$$

where $\mathbf{x}$ is a random vector of quantiles, with n independent coordinates distributed uniformly on $[0,1]$. For every $x \in [0,1)$, observe that $\gamma_\eta(x)$ and $\xi_\eta^\delta(x)$ converge towards 0 when taking η arbitrarily small, and thus $\varphi_\eta^\delta(x)$ converges towards $\varphi_0^\delta(x) := \mathbb{1}\{x \geqslant \varepsilon\} - \delta b_0^\delta(x)$. Therefore, WEV_η^δ converges towards WEV_0^δ, defined by

$$\begin{aligned}\forall \delta \in [0,1], \quad \mathrm{WEV}_0^\delta :=& \mathbb{E}_{\mathbf{x}}\left[\frac{((\mathbb{1}\{x_1 \geqslant \varepsilon\} - \delta b_0^\delta(x_1)) - (\mathbb{1}\{x_2 \geqslant \varepsilon\} - \delta b_0^\delta(x_2)))^2}{2} \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right] \\ =& \mathbb{E}_{\mathbf{x}}\left[\frac{(\varphi_0^\delta(x_1) - \varphi_0^\delta(x_2))^2}{2} \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right] = \lim_{\eta \to 0} \mathrm{WEV}_\eta^\delta.\end{aligned}$$

Proof of proposition 6.16. We are now equipped to prove the proposition. We write $\mathrm{WEV}_0^\delta = \mathbb{E}_{\mathbf{x}}\left[\varphi_0^\delta(x_1)^2 \,|\, x_1 > Y_{k+1}(\mathbf{x})\right] - \mathbb{E}_{\mathbf{x}}\left[\varphi_0^\delta(x_1)\varphi_0^\delta(x_2) \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right]$, with

$$\begin{aligned}\mathbb{E}_{\mathbf{x}}\left[\varphi_0^\delta(x_1)^2 \,|\, x_1 > Y_{k+1}(\mathbf{x})\right] &= \frac{n}{n-1}\int_0^1 \varphi_0^\delta(x)^2 G(x)\,\mathrm{d}x \\ \mathbb{E}_{\mathbf{x}}\left[\varphi_0^\delta(x_1)\varphi_0^\delta(x_2) \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right] &= \frac{n}{n-2}\int_0^1 \left(\int_t^1 \varphi_0^\delta(x)\,\mathrm{d}x\right)^2 h(t)\,\mathrm{d}t\end{aligned}$$

We next compute these quantities for uniform and discriminatory pricing. For uniform pricing ($\delta = 0$) we have that $\varphi_0^0(x) = \mathbb{1}\{x \geqslant \varepsilon\}$. We derive

$$\begin{aligned}\mathbb{E}_{\mathbf{x}}\left[\varphi_0^0(x_1)^2 \,|\, x_1 > Y_{k+1}(\mathbf{x})\right] &= \frac{n}{n-1}\int_\varepsilon^1 \widetilde{G}(x)\,\mathrm{d}x = \frac{n(1-\varepsilon)-(1-\varepsilon)^n}{n-1} \\ \mathbb{E}_{\mathbf{x}}\left[\varphi_0^0(x_1)\varphi_0^0(x_2) \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right] &= \frac{n}{n-2}\int_0^\varepsilon (1-\varepsilon)^2\widetilde{h}(t)\,\mathrm{d}t + \frac{n}{n-2}\int_\varepsilon^1 (1-t)^2\widetilde{h}(t)\,\mathrm{d}t \\ &= \frac{n(1-\varepsilon)^2 - 2(1-\varepsilon)^n}{n-2}\end{aligned}$$

and finally

$$\begin{aligned}\mathrm{WEV}_0^0 &= \frac{n(1-\varepsilon)-(1-\varepsilon)^n}{n-1} - \frac{n(1-\varepsilon)^2 - 2(1-\varepsilon)^n}{n-2} \\ &= \frac{n\left[(1-\varepsilon)^n + (1-\varepsilon)(\varepsilon(n-1)-1)\right]}{(n-1)(n-2)} \\ &\leqslant \frac{(\varepsilon n)^2/2}{n} = \frac{0.005}{n}\end{aligned}$$

For discriminatory pricing ($\delta = 1$) we have that $\varphi_0^1(x) = \mathbb{1}\{x \geqslant \varepsilon\} - b_0^1(x) = \mathbb{1}\{x \geqslant \varepsilon\}\frac{G(\varepsilon)}{G(x)}$. We will use the following bounds:

$$\begin{aligned}\int_\varepsilon^1 \frac{1}{\widetilde{G}(x)}\,\mathrm{d}x &= \int_\varepsilon^1 \frac{1}{1-(1-x)^{n-1}}\,\mathrm{d}x = \int_\varepsilon^1 \sum_{i=0}^\infty (1-x)^{(n-1)i}\,\mathrm{d}x \\ &= \sum_{i=0}^\infty \frac{(1-\varepsilon)^{(n-1)i+1}}{(n-1)i+1} \geqslant (1-\varepsilon) + \sum_{i=1}^\infty \frac{(1-\varepsilon)^{ni}}{ni} \\ &\geqslant (1-\varepsilon) + \frac{1}{n}\sum_{i=1}^\infty \frac{0.9^i}{i} = 1 - \frac{0.1}{n} - \frac{\log(0.1)}{n} \geqslant 1 + \frac{2.2}{n} \\ \int_0^1 \frac{x}{\widetilde{G}(x)}\,\mathrm{d}x &= \int_0^1 \frac{x}{1-(1-x)^{n-1}}\,\mathrm{d}x = \int_0^1 \sum_{i=0}^\infty x(1-x)^{(n-1)i}\,\mathrm{d}x \\ &= \frac{1}{2} + \sum_{i=1}^\infty \frac{1}{((n-1)i+1)((n-1)i+2)} \leqslant \frac{1}{2} + \frac{1.65}{n^2} \qquad \text{(when } n \geqslant 5\text{)}\end{aligned}$$

Next, we write

$$\mathbb{E}_{\mathbf{x}}\left[\varphi_0^1(x_1)^2 \,|\, x_1 > Y_{k+1}(\mathbf{x})\right] = \frac{n}{n-1}\int_\varepsilon^1 \frac{\widetilde{G}(\varepsilon)^2}{\widetilde{G}(x)}\,\mathrm{d}x \geqslant \frac{n\widetilde{G}(\varepsilon)^2}{n-1}\left(1+\frac{2.2}{n}\right)$$

and

$$\begin{aligned}
\mathbb{E}_{\mathbf{x}}\left[\varphi_0^1(x_1)\varphi_0^1(x_2) \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right] &= \frac{n}{n-2}\int_0^\varepsilon \left(\int_\varepsilon^1 \frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(x)}\,\mathrm{d}x\right)^2 h(t)\,\mathrm{d}t \\
&+ \frac{n}{n-2}\int_\varepsilon^1 \left(\int_t^1 \frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(x)}\,\mathrm{d}x\right)^2 h(t)\,\mathrm{d}t \\
&= \underbrace{\frac{n\widetilde{H}(\varepsilon)}{n-2}\left(\int_\varepsilon^1 \frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(x)}\,\mathrm{d}x\right)^2 + \frac{n}{n-2}\left[\left(\int_t^1 \frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(x)}\,\mathrm{d}x\right)^2 H(t)\right]_\varepsilon^1}_{=0} \\
&+ \frac{2n}{n-2}\int_\varepsilon^1 \left(\int_t^1 \frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(x)}\,\mathrm{d}x\right)\frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(t)}\widetilde{H}(t)\,\mathrm{d}t \\
&= \frac{2n}{n-2}\int_\varepsilon^1 \left(\int_\varepsilon^x \frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(t)}\widetilde{H}(t)\,\mathrm{d}t\right)\frac{\widetilde{G}(\varepsilon)}{\widetilde{G}(x)}\,\mathrm{d}x
\end{aligned}$$

Next, we will use the upper bound $\widetilde{H}(t)/\widetilde{G}(t) \leqslant 1$, which is nearly tight as $\widetilde{H}(t)/\widetilde{G}(t)$ is increasing, and has the limit $(n-2)/(n-1)$ when $t \to 0$.

$$\mathbb{E}_{\mathbf{x}}\left[\varphi_0^1(x_1)\varphi_0^1(x_2) \,|\, x_1, x_2 > Y_{k+1}(\mathbf{x})\right] \leqslant \frac{2n\widetilde{G}(\varepsilon)^2}{n-2}\int_\varepsilon^1 \frac{x}{\widetilde{G}(x)}\,\mathrm{d}x \leqslant \frac{2n\widetilde{G}(\varepsilon)^2}{n-2}\left(\frac{1}{2}+\frac{1.65}{n^2}\right)$$

Finally, we obtain

$$\begin{aligned}
\mathrm{WEV}_0^1 &\geqslant \frac{n\widetilde{G}(\varepsilon)^2}{n-1}\left(1+\frac{2.2}{n}\right) - \frac{2n\widetilde{G}(\varepsilon)^2}{n-2}\left(\frac{1}{2}+\frac{1.65}{n^2}\right) && \text{(when } n \geqslant 5\text{)} \\
&= n\widetilde{G}(\varepsilon)^2\left(\frac{2.2}{n(n-1)} - \frac{1}{(n-1)(n-2)} + \frac{3.3}{n^2(n-2)}\right) \\
&\geqslant \frac{0.01}{n} && \text{(when } n \geqslant 4\text{)}
\end{aligned}$$

$\square$

6.9.5 Proving the main theorems

Proof of lemma 6.25. Let $s_i \leqslant s_j = s_i + \varepsilon$ for some $\varepsilon > 0$. Then

$$\begin{array}{rrcl} \Leftrightarrow & u_i(s_i, \mathbf{s}_{-i}) & \leqslant & u_j(s_j, \mathbf{s}_j) \\ \Leftrightarrow & (1-c)s_i - \delta\beta^\delta(s_i) & \leqslant & (1-c)s_j - \delta\beta^\delta(s_j) \\ \Leftrightarrow & (1-c)(s_j - s_j) & \geqslant & \delta(\beta^\delta(s_j) - \beta^\delta(s_i)) \end{array}$$

Dividing by $s_j - s_i$ and taking $\varepsilon \to 0$ concludes the proof. □

Proof of proposition 6.27. First, we prove that pairwise differences is locally decreasing in δ. Let s_i, s_j with $s_i \geqslant s_j$ denote the signals of two winning bidders and $\varphi^\delta(s) := (1-c)s - \delta\beta^\delta(s)$. Note that because of lemma 6.26 (2), monotone ex-post utility holds for all $\delta \leqslant \bar{\delta}$. For all δ_1, δ_2, $0 \leqslant \delta_1 \leqslant \delta_2 \leqslant \bar{\delta}$, we have

$$|u^{\delta_1}(s_i) - u^{\delta_1}(s_j)| \geqslant |u^{\delta_2}(s_i) - u^{\delta_2}(s_j)| \tag{6.18}$$

$$\Leftrightarrow \quad |\varphi^{\delta_1}(s_i) - \varphi^{\delta_1}(s_j)| \geqslant |\varphi^{\delta_2}(s_i) - \varphi^{\delta_2}(s_i)| \tag{6.19}$$

$$\Leftrightarrow \quad -\delta_1\left(\beta^{\delta_1}(s_i) - \beta^{\delta_1}(s_j)\right) \geqslant -\delta_2\left(\beta^{\delta_2}(s_i) - \beta^{\delta_2}(s_j)\right) \tag{6.20}$$

For the final equivalence, observe that monotone ex-post utility together with lemma 6.26 (1) implies that $\frac{\delta}{1-c}\beta^\delta$ is non-expansive, allowing to remove the absolute value in eq. (6.19). lemma 6.26 (2) guarantees that eq. (6.20) holds. As the ex-post difference in utilities (eq. (6.18)) is decreasing in δ, so is its expectation. To establish global monotonicty on $[0, \bar{\delta}]$, note that if $\bar{\delta}\frac{\partial\beta^{\bar{\delta}}}{\partial s} \leqslant 1 - c$ then it also holds for any $\delta < \bar{\delta}$ by lemma 6.26 (2), concluding the proof. □

Proof of proposition 6.28. Let $u_i^\delta(s_i, \mathbf{s}_{-i})$ denote bidder i's utility in the δ-mixed auction, and let $u_i^U(s_i, \mathbf{s}_{-i})$ denote bidder i's utility in the uniform-price auction. Now let $i, j \in [n]$ be two winning bidders. As above, β^δ (resp. β^U) denotes the symmetric equilibrium bid function in the δ-mixed (resp. uniform price) auction. Let $Y_{k+1}(\boldsymbol{\beta})$ denote the first rejected bid. Then, canceling out $(1-\delta)Y_{k+1}(\boldsymbol{\beta})$, we have

$$\begin{aligned} |u_i^\delta - u_j^\delta| &= |(v_i(s_i, \mathbf{s}_{-i}) - \delta\beta^\delta(s_i)) - (v_j(s_j, \mathbf{s}_{-j}) - \delta\beta^\delta(s_j))| \\ &= |((1-c)s_i + \frac{c}{n}\sum_{k\in[n]} s_k - \delta\beta^\delta(s_i)) - ((1-c)s_j + \frac{c}{n}\sum_{k\in[n]} s_k - \delta\beta^\delta(s_j))| \\ &= |((1-c)s_i - \delta\beta^\delta(s_i)) - ((1-c)s_j - \delta\beta^\delta(s_j))| \\ &= |\varphi^\delta(s_i) - \varphi^\delta(s_j)|, \end{aligned}$$

where $\varphi^\delta(s) = (1-c)s - \delta\beta^\delta(s)$. It also holds that

$$|u_i^U - u_j^U| = |(v_i(s_i, \mathbf{s}_{-i}) - Y_{k+1}(\boldsymbol{\beta})) - (v_j(s_j, \mathbf{s}_{-j}) - Y_{k+1}(\boldsymbol{\beta}))| = |(1-c)(s_i - s_j)|.$$

We will now show that $\frac{\varphi^\delta}{1-c}$ is a non-expansive mapping. Note that φ^δ can be increasing or decreasing, so we need to show that $|\frac{\partial\varphi^\delta}{\partial s}| \leqslant 1-c$. We have $\frac{\partial\varphi^\delta}{\partial s} = 1 - c - \delta\frac{\partial\beta^\delta}{\partial s}$. As β^δ is increasing in s, $|\frac{\partial\varphi^\delta}{\partial s}| \leqslant 1-c$ holds whenever $\delta\frac{\partial\beta^\delta}{\partial s} \leqslant 2(1-c)$. Therefore

$$|u_i^\delta - u_j^\delta| = |\varphi^\delta(v_i) - \varphi^\delta(v_j)| \leqslant |(1-c)(s_i - s_j)| = |u_i^U - u_j^U| \tag{6.21}$$

Taking the square of eq. (6.21) we obtain the result point-wise, for each pair of winning signals s_i and s_j and, taking the expectation, the theorem follows. □

Theorem 6.38. *For a given common value component c, consider two δ-mixed auctions for $\delta_1 \leqslant \delta_2$ and suppose the equilibrium bidding functions β^δ satisfies $\delta_1\frac{\partial\beta^{\delta_1}(s)}{\partial s} + \delta_2\frac{\partial\beta^{\delta_2}(s)}{\partial s} \leqslant 2(1-c)$ for all signals $s \in [0, \overline{s})$. Then,* WEV *is lower for the δ_2-mixed auction than for the δ_1 one.*

Proof. Let $\varphi^\delta(s) = (1-c)s - \delta\beta^\delta(s)$. We have $u_i^\delta(\mathbf{s}) - u_j^\delta(\mathbf{s}) = \varphi^\delta(s_i) - \varphi^\delta(s_j)$. Let $\delta_1 \leqslant \delta_2$. By the generalized Cauchy mean value Theorem, we have that there exists $\xi \in [s_i, s_j]$ such that

$$|\varphi^{\delta_2}(s_i) - \varphi^{\delta_2}(s_j)| \left|\frac{\partial\varphi^{\delta_1}(\xi)}{\partial s}\right| = |\varphi^{\delta_1}(s_i) - \varphi^{\delta_1}(s_j)| \left|\frac{\partial\varphi^{\delta_2}(\xi)}{\partial s}\right|.$$

Hence if $|\frac{\partial\varphi^{\delta_2}}{\partial s}|/|\frac{\partial\varphi^{\delta_1}}{\partial s}| \leqslant 1$ then we have lower WEV for the δ_2 mixed auction. We have the following chain of equivalences:

$$\begin{aligned}
& \left|\frac{\partial\varphi^{\delta_2}(s)}{\partial s}\right| \leqslant \left|\frac{\partial\varphi^{\delta_2}(s)}{\partial s}\right|, \forall s \in (0, \overline{s}) \\
\iff\quad & \left|(1-c) - \delta_2\frac{\partial\beta^{\delta_2}(s)}{\partial s}\right| \leqslant \left|(1-c) - \delta_1\frac{\partial\beta^{\delta_1}(s)}{\partial s}\right|, \forall s \in (0, \overline{s}) \\
\iff\quad & \delta_2\frac{\partial\beta^{\delta_2}(s)}{\partial s} - (1-c) \leqslant (1-c) - \delta_1\frac{\partial\beta^{\delta_1}(s)}{\partial s}, \forall s \in (0, \overline{s}) \\
\iff\quad & \delta_1\frac{\partial\beta^{\delta_1}(s)}{\partial s} + \delta_2\frac{\partial\beta^{\delta_2}(s)}{\partial s} \leqslant 2(1-c), \forall s \in (0, \overline{s}),
\end{aligned}$$

where the third equations comes from the monotonicity of $\delta\frac{\partial\beta^\delta}{\partial s}$ in δ from lemma 6.26. □

6.9.6 Proving the bound on bid function slopes

Proof of lemma 6.31. We first rewrite $\tilde{v}(x,y)$ for $c = 1$ in terms of all the order-statistics of s_{-i}.

$$
\begin{aligned}
\tilde{v}(x,y) &= \mathbb{E}[v(s_i, s_{-i}) \mid s_i = x, Y_k = y] \\
&= \mathbb{E}[\frac{1}{n}\sum_{j\in[n]} s_j \mid s_i = x, Y_k = y] \\
&= \frac{x}{n} + \mathbb{E}[\frac{1}{n}\sum_{\substack{j\in[n],\\ j\neq i}} s_j \mid s_i = x, Y_k = y] \\
&= \frac{x}{n} + \mathbb{E}[\frac{1}{n}\sum_{j\in[n-1]} Y_j \mid s_i = x, Y_k = y] && \text{(Ordering the signals)} \\
&= \frac{x}{n} + \frac{y}{n} + \mathbb{E}[\frac{1}{n}\sum_{\substack{j\in[n-1],\\ j\neq k}} Y_j \mid s_i = x, Y_k = y] \\
&= \frac{x}{n} + \frac{y}{n} + \frac{1}{n}\sum_{j=1}^{k-1}\mathbb{E}[Y_j \mid s_i = x, Y_k = y] + \frac{1}{n}\sum_{j=k+1}^{n-1}\mathbb{E}[Y_j \mid s_i = x, Y_k = y]
\end{aligned}
$$

Note that the previous decomposition is similar to the equilibrium bid in an English auction given that k bidders have dropped out in [GO03]. However, we offer a careful derivation in the multi-unit setting of our model. We now use Theorem 2.4.1 and Theorem 2.4.2 from [ABN08] on the conditional distribution of order statistics. They state that, for $j < k$, the distribution of Y_j given $Y_k = y$ is the same as the distribution of the j-th order statistic of $k-1$ independent samples of the original distribution left-truncated at y, and we denote Z_j^l a random variable drawn according to this distribution. Hence, for $j < k$, $\mathbb{E}[Y_j \mid Y_k = y] = \mathbb{E}[Z_j^l]$. Similarly for $j > k$ we have that the distribution of Y_j given $Y_k = y$ is the same as the distribution of the $j-k$-th order statistic of $n-k-1$ independent samples of the original distribution right-truncated at y, and we denote by Z_j^r a random variable drawn according to this distribution. Hence, for $j > k$, $\mathbb{E}[Y_j \mid Y_k = y] = \mathbb{E}[Z_j^r]$. Notice that summing all order statistics drawn from some samples recovers exactly the sum of original samples. Thus we obtain

$$
\sum_{j=1}^{k-1}\mathbb{E}[Y_j \mid s_i = x, Y_k = y] = \sum_{j=1}^{k-1}\mathbb{E}[Z_j^l] = \mathbb{E}[\sum_{j=1}^{k-1} Z_j^l] = \mathbb{E}[\sum_{j=1}^{k-1} s_j \mid \forall j \in [k-1], s_j \geqslant y] = \sum_{j=1}^{k-1}\mathbb{E}[s_j \mid s_j \geqslant y].
$$

The same can be done for the Z_j^r. Finally, the s_j are iid and thus have identical conditional expectations. We obtain

$$\tilde{v}(x,y) = \frac{x}{n} + \frac{y}{n} + \frac{n-k-1}{n}\mathbb{E}[s_j \mid s_j \leqslant y] + \frac{k-1}{n}\mathbb{E}[s_j \mid s_j \geqslant y] \tag{6.22}$$

$$= \frac{x}{n} + \frac{y}{n} + \frac{n-k-1}{n}\frac{\int_0^y tf(t)\,\mathrm{d}t}{F(y)} + \frac{k-1}{n}\frac{\int_y^{\bar{s}} tf(t)\,\mathrm{d}t}{1-F(y)}, \tag{6.23}$$

which readily yields a formula for $V(s) = \tilde{v}(s,s)$. Clearly, the above function is well defined and differentiable on the open support of F.

We now examine the derivative of $V(s)$ and prove that $V'(s) \leqslant 1$. First, we consider the derivatives of the two ratios with an integral in the numerator in eq. (6.23). First, by integration by parts, we have

$$\frac{\int_0^s tf(t)\,\mathrm{d}t}{F(s)} = \frac{[tF(t)]_0^s - \int_0^s F(t)\,\mathrm{d}t}{F(s)} = s - \frac{\int_0^s F(t)\,\mathrm{d}t}{F(s)},$$

and using that for positive random variables $\int_0^{\bar{s}} tf(t)\,\mathrm{d}t = \int_0^{\bar{s}}(1-F(t))\,\mathrm{d}t = \mathbb{E}[s_i] < \infty$, which guarantees convergence of the integral, we have that

$$\begin{aligned}
\frac{\int_s^{\bar{s}} tf(t)\,\mathrm{d}t}{1-F(s)} = \frac{\mathbb{E}[s_i] - \int_0^s tf(t)\,\mathrm{d}t}{1-F(s)} &= \frac{\int_0^{\bar{s}}(1-F(t))\,\mathrm{d}t - sF(s) + \int_0^s F(t)\,\mathrm{d}t}{1-F(s)} \\
&= \frac{\int_0^{\bar{s}}(1-F(t))\,\mathrm{d}t + s(1-F(s)) - s + \int_0^s F(t)\,\mathrm{d}t}{1-F(s)} \\
&= s + \frac{\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t}{1-F(s)}
\end{aligned}$$

Now, taking derivatives, we have

$$\frac{\partial}{\partial s}\frac{\int_0^s tf(t)\,\mathrm{d}t}{F(s)} = 1 - \frac{F(s)^2 - f(s)\int_0^s F(t)\,\mathrm{d}t}{F(s)^2} = \frac{f(s)\int_0^s F(t)\,\mathrm{d}t}{F(s)^2}.$$

By a similar argument as in the proof of lemma 6.32, using log-concavity of f, the above derivative is bounded by 1. Taking the derivative of the second ratio, we have

$$\frac{\partial}{\partial s}\left(s + \frac{\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t}{1-F(s)}\right) = 1 + \frac{-(1-F(s))^2 + f(s)\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t}{(1-F(s))^2} = \frac{f(s)\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t}{(1-F(s))^2}. \tag{6.24}$$

To derivative of $\log(\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t)$:

$$\frac{\partial^2}{(\partial s)^2}\log(\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t) = \frac{\partial}{\partial s}\frac{-(1-F(s))}{\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t} = \frac{f(s)\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t - (1-F(s))^2}{\left(\int_s^{\bar{s}}(1-F(t))\,\mathrm{d}t\right)^2}. \tag{6.25}$$

eq. (6.25) is negative iff $f(s)\int_s^{\overline{s}}(1-F(t))\,\mathrm{d}t/(1-F(s))^2 \leqslant 1$. This means that the log-concavity of $\int_s^{\overline{s}}(1-F(t))\,\mathrm{d}t$ is equivalent to eq. (6.24) being smaller than 1. As the log-concavity of $\int_s^{\overline{s}}(1-F(t))\,\mathrm{d}t$ follows from the log-concavity of f and $(1-F)$ [BB05, Theorem 3], $f(s)\int_s^{\overline{s}}(1-F(t))\,\mathrm{d}t/(1-F(s))^2 \leqslant 1$ is implied. Finally, using the above derivatives it is clear that $V'(s) > 0$, and

$$V'(s) \leqslant \frac{2}{n} + \frac{n-k-1}{n} \cdot 1 + \frac{k-1}{n} \cdot 1 = 1.$$

□

Proof of lemma 6.32. Let us compute the second derivative of the logarithm of $\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y$:

$$\begin{aligned}
\frac{\partial^2 \log\left(\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y\right)}{(\partial s)^2} &= \frac{\partial}{\partial s}\left(\frac{G^{\frac{1}{\delta}}(s)}{\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y}\right) \\
&= \frac{\frac{1}{\delta}g(s)G^{\frac{1}{\delta}-1}(s)\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y - G^{\frac{2}{\delta}}(s)}{\left(\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y\right)^2} \\
&= \frac{G^{\frac{1}{\delta}-1}(s)}{\left(\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y\right)^2}\left(\frac{1}{\delta}g(s)\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y - G^{\frac{1}{\delta}+1}(s)\right).
\end{aligned}$$

Notice that the left-hand fraction is always positive. Hence log-concavity of $\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y$ is equivalent to $\frac{1}{\delta}g(s)\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y - G^{\frac{1}{\delta}+1}(s)$ being negative. The latter is equivalent to

$$1 \geqslant \frac{g(s)\int_0^s G^{\frac{1}{\delta}}(y)\,\mathrm{d}y}{\delta G^{\frac{1}{\delta}+1}(s)} = \frac{\partial \beta^\delta(s)}{\partial s}.$$

□

Corollary 6.39. *For uniformly distributed signals, any pricing dominant in pairwise differences contains a discriminatory proportion of at least* $\frac{2n(1-c)}{2n-c(n-2)}$*, and for exponentially distributed signals at least* $\frac{2n(1-c)}{2n-c(n-(k+1))}$.

Proof of corollary 6.39. While $\sup_{[0,\overline{s})}\frac{\partial\beta^\delta}{\partial s}$ can be difficult to compute analytically even for simple distributions, it is sometimes possible to compute $\sup_{[0,\overline{s})}V'(s)$. For the uniform distribution, we have $\sup_{[0,\overline{s})}V'(s) = 1 - c\frac{n-2}{2n}$. Thus, using the same argument as in the proof of theorem 6.18, it follows that $\delta^*(c) \geqslant \frac{2n(1-c)}{2n-c(n-2)} \to_{n\to\infty} \frac{(1-c)}{1-c/2}$. For the exponential distribution, we have $\sup_{[0,\overline{s})}V'(s) = 1 - c(\frac{1}{2} - \frac{k+1}{2n})$, and thus $\delta^*(c) \geqslant \frac{2n(1-c)}{2n-c(n-(k+1))} \to_{n\to\infty} \frac{(1-c)}{1-c/2}$. □

Proof of lemma 6.33. To prove lemma 6.33, we will use properties of log-concave distributions from [BB05]. Namely their Theorems 1 and 3 state together that log-concavity of a density f implies log-concavity of the corresponding cdf F and of the complementary cdf $1 - F$, and that log-concavity of F or $1 - F$ imply log-concavity of respectively $\int_0^s F$ or $\int_s^{\overline{s}} F$, where $\overline{s}$ is the upper limit of the support of f (either a constant or $+\infty$). Additionally, we also have that the product of two log-concave functions is log-concave also. Using the above properties, we have that F and $1 - F$ are log-concave.

Moreover, alternative expression for the order statistics are given, e.g., in [Fis65].

$$G_m^n(s) = \frac{n!}{(n-m)!(m-1)!} \int_0^{F(s)} t^{n-m}(1-t)^{m-1} dt$$

and

$$g_m^n(s) = \frac{n!}{(n-m)!(m-1)!} F(s)^{n-m}(1-F(s))^{m-1} f(s). \tag{6.26}$$

Thus, the order statistics density g, given by eq. (6.26), is a product of F, $1 - F$, and f, and g as well as the corresponding cdf G are also log-concave. Furthermore, $G^{\frac{1}{\delta}}$ is log-concave because $\log(G^{\frac{1}{\delta}}) = \delta \log(G)$. Finally, we remark that $G^{\frac{1}{\delta}}$ is right-continuous non-decreasing by composition with $x \mapsto x^{\frac{1}{\delta}}$, which is continuous non-decreasing, and $G^{\frac{1}{\delta}}(0) = 0$, as well as $G^{\frac{1}{\delta}}(\overline{s}) = 1$ (if $\overline{s} = \infty$, the equality is understood as a limit). Therefore $G^{\frac{1}{\delta}}$ is a cdf, and applying one last time [BB05], we obtain that $\int_0^s G^{\frac{1}{\delta}}$ is log-concave. $\square$

6.9.7 Numerical experiments

Lemma 6.40. *Suppose an auction is a* winners-pay auction. *Then we can write* $E_{\mathbf{s}}[u_1 \ |1 \ wins] = \frac{n}{k} E_{\mathbf{s}}[u_1]$, $E_{\mathbf{s}}[u_1^2|1 \ wins] = \frac{n}{k} E_{\mathbf{s}}[u_1^2]$, *and* $E_{\mathbf{s}}[u_1 u_2|1 \ and \ 2 \ win] = \frac{n(n-1)}{k(k-1)} E_{\mathbf{s}}[u_1 u_2]$.

Proof. Observe that we have

$$\mathbb{E}_s[u_1^2 \mid 1 \text{ and } 2 \text{ win}] = \mathbb{E}_s[u_1^2 \mid 1 \text{ wins}] = \frac{\mathbb{E}[u_1^2]}{\mathbb{P}[1 \text{ wins}]} = \frac{n}{k} \cdot \mathbb{E}[u_1^2] \tag{6.27}$$

$$\mathbb{E}_s[u_1 u_2 \mid 1 \text{ and } 2 \text{ win}] = \frac{\mathbb{E}[u_1 u_2]}{\mathbb{P}[1 \text{ and } 2 \text{ win}]} = \frac{n(n-1)}{k(k-1)} \cdot \mathbb{E}[u_1 u_2] \tag{6.28}$$

$\square$

6.9.8 Discussion

Proof of proposition 6.20. We define the probability that i wins $q_i(s_i) := \mathbb{P}_{\mathbf{s}_{-i}}[i \text{ wins}]$. Recall that $b^D(s_i)$ denotes the equilibrium bid in the pay-as-bid auction. Consider any standard auction, characterised by a payment rule $(p_1(s), \dots, p_n(s))$. Revenue equivalence implies that

$$q_i(s_i) \cdot b^D(s_i) = \mathbb{E}_{\mathbf{s}_{-i}}[b^D(s_i) \cdot \mathbb{1}\{i \text{ wins}\}] = \mathbb{E}_{\mathbf{s}_{-i}}[p_i(s)]. \tag{6.29}$$

In particular, note that if p_i is chosen to be the uniform pricing rule, this formula can be used to compute $b^D(s_i)$. Now define the ex-post surplus $u_i(s_i, \mathbf{s}_{-i}) := v(s_i) \cdot \mathbb{1}\{i \text{ wins}\} - p_i(s)$. We write

$$u_i(s_i, \mathbf{s}_{-i}) = \underbrace{\mathbb{1}\{i \text{ wins}\} \cdot (v(s_i) - b^D(s_i))}_{u_i^D(s)} + \underbrace{\mathbb{1}\{i \text{ wins}\} \cdot b^D(s_i) - p_i(s)}_{\delta(s)}. \tag{6.30}$$

Now, observe that by revenue equivalence we have $\mathbb{E}_{\mathbf{s}_{-i}}[\delta(s_i, \mathbf{s}_{-i})] = 0$ for all s_i. We write

$$\mathbb{E}_{\mathbf{s}_{-i}}[u_i(s_i, \mathbf{s}_{-i})]^2 = \mathbb{E}_{\mathbf{s}_{-i}}[u_i^D(s_i, \mathbf{s}_{-i})]^2 \tag{6.31}$$

$$\mathbb{E}_{\mathbf{s}_{-i}}[u_i(s_i, \mathbf{s}_{-i})^2] = \mathbb{E}_{\mathbf{s}_{-i}}[u_i^D(s_i, \mathbf{s}_{-i})^2] + 2\underbrace{\mathbb{E}_{\mathbf{s}_{-i}}[u_i^D(s_i, \mathbf{s}_{-i}) \cdot \delta(s_i, \mathbf{s}_{-i})]}_{\geqslant 0} + \underbrace{\mathbb{E}_{\mathbf{s}_{-i}}[\delta(s_i, \mathbf{s}_{-i})^2]}_{\geqslant 0} \tag{6.32}$$

To show that the extra terms are non-negative, notice that $\delta(s_i, \mathbf{s}_{-i})^2 \geqslant 0$, and that

$$\mathbb{E}_{\mathbf{s}_{-i}}[u_i^D(s_i, \mathbf{s}_{-i}) \cdot \delta(s_i, \mathbf{s}_{-i})] = \underbrace{(v(s_i) - b^D(s_i))}_{\geqslant 0} \cdot \underbrace{(q_i(s_i) \cdot b^D(s_i) - \mathbb{E}_{\mathbf{s}_{-i}}[\mathbb{1}\{i \text{ wins}\} \cdot p_i(s)])}_{\geqslant \mathbb{E}[\delta(s)] = 0} \tag{6.33}$$

Therefore, putting everything together, we obtain

$$\begin{aligned} \mathrm{Var}_{\mathbf{s}_{-i}}[u_i(s_i, \mathbf{s}_{-i})] &= \mathbb{E}_{\mathbf{s}_{-i}}[u_i(s_i, \mathbf{s}_{-i})^2] - \mathbb{E}_{\mathbf{s}_{-i}}[u_i(s_i, \mathbf{s}_{-i})]^2 && (6.34) \\ &\geqslant \mathbb{E}_{\mathbf{s}_{-i}}[u_i^D(s_i, \mathbf{s}_{-i})^2] - \mathbb{E}_{\mathbf{s}_{-i}}[u_i^D(s_i, \mathbf{s}_{-i})]^2 && (6.35) \\ &= \mathrm{Var}_{\mathbf{s}_{-i}}[u_i^D(s_i, \mathbf{s}_{-i})] && (6.36) \end{aligned}$$

Finally, observe that an auction which minimize the interim variance also minimize the ex-ante variance. Denoting by u_i the utility of a bidder in the pay-as-bid auction, the law of total variance states

$$\mathrm{Var}_s[u_i] = \mathbb{E}_{s_i}[\mathrm{Var}_{\mathbf{s}_{-i}}[u_i(s_i, \mathbf{s}_{-i})]] + \mathrm{Var}_{s_i}[\mathbb{E}_{\mathbf{s}_{-i}}[u_i]]. \tag{6.37}$$

By revenue equivalence, we know that $\mathbb{E}_{\mathbf{s}_{-i}}[u_i]$ is the same for all standard auctions, hence $\mathrm{Var}_{s_i}[\mathbb{E}_{\mathbf{s}_{-i}}[u_i]]$ is also the same for standard auctions (it only depends on the signal distribution).

The interim variance $\mathrm{Var}_{\mathbf{s}_{-i}}[u_i(s_i, \mathbf{s}_{-i})]$ is minimal point-wise (in s_i) for all standard auctions, hence is also minimal in expectation. Therefore, the ex-ante variance is minimal in the pay-as-bid auction among standard auctions. □

Proof of proposition 6.22. We first derive the equilibrium bid for the first-rejected-bid uniform auction with common values and reserve price $r > 0$. Fix a signal s. Let β be an increasing symmetric equilibrium, and let $s_r = \inf\{s \geqslant 0 \mid \beta(s_r) \geqslant r\}$ be the threshold signal for the bid to exceed a given reserve price r. For $z \geqslant s_r$, we consider $U(s_i, z)$, the expected payoff of bidding $\beta(z)$ with signal s_i:

$$\begin{aligned}U(s_i, z) &= \int_0^z \widetilde{V}(s_i, y)g(y)\,\mathrm{d}y - \int_0^{s_r} rg(y)\,\mathrm{d}y - \int_{s_r}^z \beta(y)g(y)\,\mathrm{d}y \\ &= \int_0^z \widetilde{V}(s_i, y)g(y)\,\mathrm{d}y - rG(s_r) - \int_{s_r}^z \beta(y)g(y)\,\mathrm{d}y.\end{aligned}$$

If $z < s_r$ then the bid is below the reserve price, no item is won, and $U(s_i, z) = 0$. If the payoff is maximized for $z \geqslant s_r$, then, by solving the first order condition, a bid of $\widetilde{V}(s_i, s_i) = V(s_i)$ is optimal. Hence, bidding $V(s_i)$ is preferred to bidding zero if the expected payoff is greater than zero. Because $V(s_i)$ is increasing and continuous, these two payoffs are equal for $s_i = s_r$ by definition: s_r corresponds to the threshold signal beyond which a positive bid of $V(s_i)$ is preferred to a zero profit. The equation

$$U(s_r, s_r) = \int_0^{s_r} \widetilde{V}(s_r, y)g(y) - rG(s_r) = 0, \tag{6.38}$$

implicitly characterizes s_r. The equilibrium bid is $\beta_r^{\delta=0} = V(s_i)$ for $s_i \geqslant s_r$ and $\beta_r^{\delta=0} = 0$ otherwise.

Using revenue equivalence, we derive the equilibrium bid in the pay-as-bid auction. We have that, for $s_i \geqslant s_r$,

$$\beta_r^{\delta=1}(s_i) = \int_0^{s_r} \frac{rg(y)}{G(s_i)}\,\mathrm{d}y + \int_{s_r}^{s_i} \frac{V(y)g(y)}{G(s_i)}\,\mathrm{d}y = V(s_i) + (r - V(s_r))\frac{G(s_r)}{G(s_i)} - \int_{s_r}^{s_i} \frac{V'(y)G(y)}{G(s_i)}\,\mathrm{d}y. \tag{6.39}$$

Taking the derivative yields

$$\frac{\partial \beta_r^{\delta=1}(s_i)}{\partial s_i} = \frac{g(s_i)}{G^2(s_i)}\left((V(s_r) - r)G(s_r) + \int_{s_r}^{s_i} V'(y)G(y)\,\mathrm{d}y\right)$$

$$
\begin{aligned}
&= \frac{g(s_i)}{G^2(s_i)} \left(\int_0^{s_r} V(y)g(y)\,\mathrm{d}y - rG(S_r) + \int_0^{s_i} V'(y)G(y)\,\mathrm{d}y \right) \\
&= \frac{g(s_i)}{G^2(s_i)} \left(\int_0^{s_r} V(y)g(y)\,\mathrm{d}y - rG(S_r) + \int_0^{s_i} V'(y)G(y)\,\mathrm{d}y \right) \\
&= \frac{g(s_i)}{G^2(s_i)} \left(\int_0^{s_r} (\widetilde{V}(y,y) - \widetilde{V}(s_r,y))g(y) + \int_0^{s_i} V'(y)G(y)\,\mathrm{d}y \right) \\
&\leqslant \frac{g(s_i)}{G^2(s_i)} \int_0^{s_i} V'(y)G(y)\,\mathrm{d}y \\
&= \frac{\partial \beta_{r=0}^{\delta=1}(s_i)}{\partial s}.
\end{aligned}
$$

We use eq. (6.38) for the second-to-last equality, and the fact that $\widetilde{V}(y,y) \leqslant \widetilde{V}(s_r,y)$ for $y \leqslant s_r$, by monotonicity of $\widetilde{V}$, for the inequality. $\square$

7 The price of opportunity fairness in matroid allocation problems

Abstract. We consider matroid allocation problems under *opportunity fairness* constraints: resources need to be allocated to a set of agents under matroid constraints (which includes classical problems such as bipartite matching). Agents are divided into C groups according to a sensitive attribute, and an allocation is opportunity-fair if each group receives the same share proportional to the maximum feasible allocation it could achieve in isolation. We study the Price of Fairness (PoF), i.e., the ratio between maximum size allocations and maximum size opportunity-fair allocations. We first provide a characterization of the PoF leveraging the underlying polymatroid structure of the allocation problem. Based on this characterization, we prove bounds on the PoF in various settings from fully adversarial (worst-case) to fully random. Notably, one of our main results considers an arbitrary matroid structure with agents randomly divided into groups. In this setting, we prove a PoF bound as a function of the (relative) size of the largest group. Our result implies that, as long as there is no dominant group (i.e., the largest group is not too large), opportunity fairness constraints do not induce any loss of social welfare (defined as the allocation size). Overall, our results give insights into which aspects of the problem's structure affect the trade-off between opportunity fairness and social welfare.

Contents

7.1 Introduction

Allocating scarce resources among agents is a fundamental task in diverse fields such as online markets [CS98], online advertising [Meh13], the labor market [Com23], university admissions [Akb+22; GS13], refugee programs [Aha+21; DKT23; Fre+23], or organ transplants [Akb+20]. Traditionally, central planners aim to efficiently compute allocations that maximize some metric of social welfare such as the total number of allocated resources. Unfortunately, optimal allocations that neglect equity considerations often result in disparate treatment and unfair outcomes for legally protected groups of individuals, as documented in domains such as job offerings [LT19a; Spe+18b] and online advertising [Ali+19b; BG18b], among others.

Matching markets, a prominent instance of resource allocation problems, are especially sensitive to these discrimination concerns. For example, initiatives such as the European Union's proposed job-matching platform for migrants [Com23], and the urgent demands arising from the global refugee crisis [Ref23] are complex challenges where fairness must be accounted for: migrants can belong to different demographic groups defined by sensitive attributes such as age, ethnicity, gender, or wealth, and jobs or resettlement locations must be allocated in a fair manner.

Motivated by these challenges, we consider *matroid allocation problems*, i.e., resource allocation problems where the constraint has a matroid structure [Sch+03, Chapter 44]. An instance of our problem is defined by a pair $(E, \mathcal{I})$ where E is a finite set of agents and $\mathcal{I}$ the set of feasible resource allocations, and such that E is partitioned based on sensitive attributes into $C \in \mathbb{N}$ distinct groups (for the formal definitions, please refer to Section 7.2). Matroid allocation problems have a theoretical structure that gives tractability while being expressive enough to formulate many important problems. They include our main application of bipartite matching discussed above, but also other allocation problems where fairness is relevant. For instance, the maintenance of communication networks between different cities, where the state's goal is to fairly distribute the maintenance rights among diverse companies to ensure healthy competition; or the selection of members for a constitutional commission that must satisfy parity requirements such as gender and ethnic representation; can both be formulated as matroid allocation problems.

Prior works have tackled fairness challenges in (matroid) allocation problems. [Chi+19] study fair matroid allocation problems, and [Ban+23] examine fair matching (see a more complete discussion in Section 7.5). These works, however, focus on the efficient computation of fair allocations and provide approximation algorithms.

In contrast, we focus on the *quality* of the optimal fair allocation. Indeed, imposing fairness constraints may reduce social welfare (the total number of resources allocated). To understand the trade-off between fairness and social welfare, we study the metric called *Price of Fairness*, introduced independently by [BFT11] and [Car+12] in the contexts of proportional fairness and equitability, respectively. This metric, formally defined by

$$\mathrm{PoF}(\mathcal{I}) := \frac{\max\{|S| : S \in \mathcal{I}\}}{\max\{|S| : S \in \mathcal{I} \text{ is fair}\}},$$

where $|S|$ denotes the size of the allocation S, provides insights into the scenarios where fairness leads to a degradation of the social welfare.

We focus on a novel notion of fairness that we introduce: **opportunity fairness**. Opportunity fairness draws inspiration from the notion of Equality of Opportunity [Har+16c] in machine learning and from Kalai-Smorodinsky fairness [KS75] in fair division. This fairness notion is particularly adapted to the structure of the allocation problem as it accounts for the inherent capabilities of the agents groups. Formally, an allocation is called opportunity fair if for any two groups c, c', it satisfies

$$\frac{\text{\# Resources allocated to } c}{\text{Total \# of resources that can be allocated to } c} = \frac{\text{\# Resources allocated to } c'}{\text{Total \# of resources that can be allocated to } c'}.$$

We consider a **large market setting** where the number of resources that can be allocated to each group is large, as is often the case, e.g., in job market platforms. In this situation, integral allocations can be well approximated by fractional—or randomized—allocations (as proved in Section 7.2.4). Hence, we focus on the PoF computation under *fractional allocations*.

Contributions. The price of opportunity fairness may depend on different features of the problem, in particular the set of feasible allocation and the agents' group assignment. We prove tight PoF bounds in multiple settings from adversarial to fully random, in line with the beyond-worst-case paradigm [Rou20], which seeks to capture more realistic and nuanced behaviors than worst-case analysis alone:

1. *Polymatroid representation:* We first show that for matroids constraints, the set of feasible per-group allocations can be represented as a *polymatroid*, a multi-set generalization of matroids. We then leverage the polymatroid representation to achieve a simpler characterization for the price of opportunity fairness, which is key for the subsequent analysis (Proposition 7.7).

2. *Adversarial analysis:* When both the group partition and the set of feasible allocations are chosen adversarially, we show that the worst-case PoF is $C-1$, regardless of the number of agents (Theorem 7.9).

3. *Parametrized families of matroids:* By considering a parametrized family of matroids, we conduct a finer PoF analysis with bounds that interpolate the best case of no loss, $\mathrm{PoF} = 1$, and the worst possible case, $\mathrm{PoF} = C - 1$ (Propositions 7.12 and 7.14).

4. *Semi-random setting:* When agents are randomly partitioned into C groups according to some distribution $p = (p_1, \cdots, p_C)$, we characterize the worst-case PoF as a function of $\max_{c \in [C]} p_c$ by reducing a joint infinite-dimensional combinatorial optimization problem to a one-dimensional optimization problem (Theorem 7.16). Remarkably, we show that as long as $\max_{c \in [C]} p_c \leqslant 1/(C-1)$, no social welfare loss is incurred under opportunity fairness constraints, in particular suggesting that *the trade-off between fairness and social welfare is not entirely due to the presence of small groups, but rather due to the presence of a dominant group* (Corollary 7.17).

5. *Random graphs model:* Finally, we extend the no-social-welfare-loss result of the previous case to any groups partition distribution p whenever the set of feasible allocations $\mathcal{I}$ is obtained from certain Erdös-Rényi random graphs (Propositions 7.18 and 7.19).

Overall, our results lead to a better understanding of how the structure of both the agents groups and feasible allocations can affect the price of fairness. The main qualitative takeaway is that *for realistic matroid and protected groups instances, opportunity fair allocations incur only a small social welfare loss.* While our main focus is on opportunity fairness, as it is the most relevant fairness notion for the setting we consider, we also provide additional results for other fairness notions in Section 7.10.

7.2 Model

7.2.1 Matroids and Colored Matroids

Let E be a finite set of agents. We denote by $\mathcal{I} \subseteq 2^E$ a family of **feasible allocations**, where for any allocation $S \in \mathcal{I}$, $e \in S$ represents that agent e got a resource allocated. We assume that $(E, \mathcal{I})$ is a finite matroid:

Definition 7.1. The pair $(E, \mathcal{I})$ is a (finite) **matroid** if $\emptyset \in \mathcal{I}$ and the following properties are satisfied: (i) For any $S \in \mathcal{I}$ and $T \subseteq S$, $T \in \mathcal{I}$ (*hereditary property*); and (ii) For any $S, T \in \mathcal{I}$, such that $|S| < |T|$, there exists $e \in T \setminus S$ such that $S \cup \{e\} \in \mathcal{I}$ (*augmentation property*).

Matroids are particularly useful for combinatorial optimization. They are rich enough to describe many allocation problems often encountered in practice, e.g., the bipartite matching (traversal matroids), the communication network (graphic matroids), and the constitutional commission problem (uniform matroid) mentioned in the introduction (see Section 7.6 for the definition of each sub-class of matroids). More importantly, due to the augmentation property, a maximal size allocation under matroid constraints can be computed in polynomial time via a greedy algorithm, provided there is a polynomial-time oracle to identify if a set is feasible.

To every matroid $(E, \mathcal{I})$, we associate a **rank function** $\mathrm{r} : 2^E \to \mathbb{R}_+$, which maps each $S \subseteq E$ to $\mathrm{r}(S) := \max\{|T| \mid T \subseteq S, T \in \mathcal{I}\}$, that is, to the size of the maximum feasible allocation included in S. Basic results of matroid theory [Sch+03] show that the rank function is submodular[1], non-decreasing, and that $0 \leqslant \mathrm{r}(S) \leqslant |S|$ for any $S \subseteq E$.

We consider matroids in which the set of agents is partitioned into C groups (based on sensitive attributes)—also termed colors [Chi+19]. Given a matroid $(E, \mathcal{I})$, $C \in \mathbb{N}$, and $(E_c)_{c\in[C]}$ a partition of E into groups, the tuple $((E_c)_{c\in[C]}, \mathcal{I})$ is called a ***C*-colored matroid** (or simply colored matroid).

Given a colored matroid and a subset of groups $\Lambda \subseteq [C]$, we denote by $\mathrm{r}(\Lambda) := \mathrm{r}(\cup_{c\in\Lambda} E_c)$ the rank of the corresponding subset of agents. We call the function $\mathrm{r} : 2^{[C]} \to \mathbb{R}_+$ the rank function of the colored matroid. The rank function of the colored matroid inherits all properties from the rank function of the original matroid. In addition, remark that $\mathrm{r}([C])$ corresponds to the size of a maximum size allocation within $\mathcal{I}$, i.e., the maximum social welfare achievable in the corresponding resource allocation problem, while $\mathrm{r}(c)$[2] corresponds to the *opportunity level* of the color (the group) c, i.e., the maximum social welfare when considering only the agents within E_c.

[1] A set function f is submodular if for all finite sets S and T, $f(S) + f(T) \geqslant f(S \cap T) + f(S \cup T)$.
[2] We write $\mathrm{r}(c)$ instead of $\mathrm{r}(\{c\})$ for convenience, since there is no ambiguity.

7.2.2 Opportunity Fairness

Given a colored matroid $((E_c)_{c\in[C]}, \mathcal{I})$, we denote $\mathrm{I} := \{x \in \mathbb{N}^C \mid$ there exists $S \in \mathcal{I}, x_c = |S \cap E_c|$, for any $c \in [C]\}$. The set I of **integer feasible group allocations** will be sufficient to find optimal fair allocations. We consider fractional allocations as feasible solutions (see Section 7.2.4 for the justification): we denote by $M :=$ convex hull(I) the set of **fractional feasible group allocations**.

With this notation, we introduce opportunity fairness and the price of fairness for fractional allocations:

Definition 7.2. A fractional allocation $x \in M$ is **opportunity fair** if for any $c, c' \in [C]$,

$$\frac{x_c}{\mathrm{r}(c)} = \frac{x_{c'}}{\mathrm{r}(c')}.$$

We denote by F the **set of opportunity fair fractional allocations**. Then the **Price of Opportunity Fairness** (PoF) is defined as

$$\mathrm{PoF}(M) := \frac{\max_{x\in M} \sum_{c\in[C]} x_c}{\max_{x\in F} \sum_{c\in[C]} x_c}.$$

7.2.3 Comparison with proportional fairness and other fairness notions

Proportional fairness, a broadly studied fairness notion in the literature, aims to equalize the ratios $x_c/|E_c|$, i.e., the number of allocated resources to each group relative to their size. In machine learning, proportional fairness corresponds to *demographic parity* [BHN23] while *opportunity fairness* is more closely related to *equality of opportunity*, as it accounts for the inherent quality of the groups.

A key limitation of proportional fairness in the matroid allocation setting is that it is *sensitive to the presence of irrelevant agents*, unlike opportunity fairness. Indeed, whenever the allocation problem possesses agents that cannot be allocated any resources, proportional fairness becomes too constraining to satisfy. For instance, adding irrelevant agents of color c increases $|E_c|$ without increasing the size of the feasible allocations to that group, thereby reducing the ratio $x_c/|E_c|$. To maintain proportional fairness, the allocation to other groups must be reduced, even though the underlying allocation problem remains unchanged.

This highlights a pathological behavior: even under fractional allocations, the *Price of Proportional Fairness* can be unbounded as the addition of irrelevant agents drives all fair allocations of other groups to zero. In contrast, as we will show throughout the paper, *opportunity fairness is robust to such manipulations* and remains always bounded, as it takes into account the structure of the allocation problem. In Section 7.10, we provide a detailed discussion of the relationships between different notions of fairness (including weighted fairness and leximin fairness), along with their respective prices of fairness; which further highlights the relevance of opportunity fairness in our context.

7.2.4 From Fractional Allocations to Approximately Fair Integral Allocations

We consider the relaxation to fractional allocations for two main reasons. First, when restricted to integer allocations, many resource allocation problems have no opportunity fair allocations besides the empty one. Indeed, even for two colors, whenever $\mathrm{r}(1)$ and $\mathrm{r}(2)$ are co-prime and $(\mathrm{r}(1), \mathrm{r}(2)) \notin \mathrm{I}$, the only feasible fair integral allocation is to allocate 0 resources to each group (see Section 7.7).

Second, by slightly relaxing the fairness constraints, any optimal fair fractional allocation can be implemented in an integral fashion through a specific randomized rounding technique, at a cost that becomes negligibly small in large markets. We first define the relaxed fairness notion:

Definition 7.3. For $\gamma \in [0,1]$, an allocation $x \in M$ is said to be **γ-opportunity fair** if for any $c, c' \in [C]$, it holds that $x_c/\mathrm{r}(c) \geqslant \gamma \cdot x_{c'}/\mathrm{r}(c')$.

Setting $\gamma = 1$ recovers Definition 7.2, while $\gamma = 0$ corresponds to no fairness considerations.

Proposition 7.4. *For any fractional opportunity fair maximum size allocation $x \in F$, there exists a random allocation X, such that, $\mathbb{E}[X] = x$ and for all realizations, X is feasible, integral, $\left(1 - \frac{2C}{\min_c r(c)}\right)$-opportunity fair, and $\|X - x\|_1 \leqslant C$.*

Proof Sketch. The argument relies on the polymatroid characterization of M (Proposition 7.7) that will be proven independently in Section 7.3. Then by Theorem 35 of [Edm70], which guarantees that the extreme points of the intersection of two integral polymatroids remain integral, x can be written as a convex combination of nearby integral allocations. See Section 7.9.1 for the proof. $\square$

In a **large market**, that is, when $\min_{c\in[C]} r_c$ is large, Proposition 7.4 shows that both the gap between $\mathrm{PoF} = r([C])/\|x\|_1$ and $r([C])/\|X\|_1$, and the fairness degradation, become negligible. Hence, it provides the strongest fairness guarantees ex-ante (since $\mathbb{E}[X] = x$ and x is perfectly fair), jointly with an approximately fair integral allocation at low cost ex-post. Such a best-of-both-worlds approach that leverages fractional allocations to design lotteries satisfying ex-ante requirements (here fairness) and additional ex-post properties is widespread in market design [AN19; BM01b; Bud+13; HZ79] and fair division [Azi20; FSV20], with applications in school choice, housing, and kidney exchange, to name but a few.

Even when considering fractional allocations, one may want to impose approximate fairness constraints only. Interestingly, for any colored matroid, our bounds on PoF (with exact fairness constraint from Definition 7.2) can easily be transposed into bounds for the relaxed setting. Formally, denote by PoF_γ the price of γ-opportunity fairness (i.e., when fairness is defined by Definition 7.3).

Proposition 7.5. *Let M be a C-colored matroid. Then, either $\gamma < \max_{x\in\mathrm{F}} \min_{c\in[C]} x_c/r_c$ and $\mathrm{PoF}_\gamma(M) = 1$, or*

$$\gamma \cdot \mathrm{PoF}(M) \leqslant \mathrm{PoF}_\gamma(M) \leqslant \frac{\gamma \cdot \mathrm{PoF}(M)}{1 - (1-\gamma)^{\frac{((C-1)-\mathrm{PoF}(M))}{C-2}}}.$$

See proof in Section 7.9.2. Proposition 7.5 allows us to seamlessly lift any PoF upper and lower bounds to PoF_γ. As before, the same randomized rounding applies, at a small cost. Thus, **the rest of the paper will focus on fractional allocation and perfect fairness constraints**, as Propositions 7.4 and 7.5 can handle integrality and relaxed constraint considerations.

7.3 Polymatroid Structure and PoF Characterization

This section is devoted to characterizing the price of opportunity fairness of a matroid as a simple combinatorial optimization problem. Our main technique will be the use of polymatroids.

Definition 7.6. The **polymatroid** associated to the submodular function $f : 2^C \to \mathbb{R}_+^C$ is the polytope

$$\{x \in \mathbb{R}_+^C \mid \textstyle\sum_{c\in\Lambda} x_c \leqslant f(\Lambda), \forall \Lambda \subseteq [C]\}.$$

Polymatroids can be seen as a generalization of matroids, as there is a natural mapping from a matroid on a ground set E to a polymatroid included in $[0,1]^E$, where a feasible allocation $S \in \mathcal{I}$ is associated to a vector $z \in [0,1]^E$ with coordinates $z_e = 1$ if $e \in S$ and 0 otherwise. Polymatroids are strictly more general, as coordinates can be larger than 1. Proposition 7.7 shows that there is also a natural relation between colored matroids and polymatroids. Refer to Section 7.9.3 for the proof.

Proposition 7.7. *Let $((E_c)_{c\in[C]}, \mathcal{I})$ be a colored matroid with rank function r and set of feasible fractional allocations M. Then, M is the polymatroid associated to the function r, i.e.,*

$$M = \{x \in \mathbb{R}_+^C \mid \sum_{c\in\Lambda} x_c \leqslant \mathrm{r}(\Lambda), \forall \Lambda \subseteq [C]\}.$$

Note that while the usual natural mapping from matroids to polymatroids is not a surjection (the coordinates must remain bounded by 1, which is not the case of all polymatroids), the mapping $((E_c)_{c\in[C]}, \mathcal{I}) \mapsto M$ from the set of colored matroids to the set of all Polymatroids is a surjection. From now on, we will use interchangeably the names feasible fractional allocations, polymatroid, and colored matroid for M.

The set M inherits interesting properties from being a polymatroid. For instance, the Pareto frontier of the Multi-objective Optimization Problem $\max_{x\in M}(x_1, x_2, ..., x_C)$ corresponds to the set of allocations maximizing social welfare $\sum_{c\in[C]} x_c$ [HH02]. In particular, the existence of an allocation of maximum size that is opportunity fair is reduced to verifying whether the intersection between the Pareto frontier and the line defined by the opportunity fairness condition is non-empty. Figure 7.1 illustrates M and P for a C-colored matroid with $C = 2$ and $C = 3$, respectively. Remark that P is simultaneously the Pareto frontier and the set of points which maximize $\sum_{c\in[C]} x_c$.

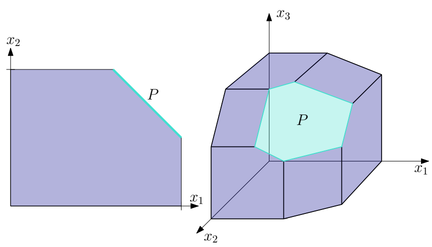


Fig. 7.1: Examples of the set of fractional feasible allocations M (dark blue solid region) and the Pareto frontier P (light blue region) for $C = 2$ and $C = 3$.

The structure of M yields the following characterization of PoF. Refer to Section 7.9.4 for the proof.

Corollary 7.8. *The price of opportunity fairness of a polymatroid M is given by,*

$$\mathrm{PoF}(M) = \frac{\mathrm{r}([C])}{\sum_{c\in[C]} \mathrm{r}(c)} \cdot \max_{\Lambda\subseteq[C]} \frac{\sum_{c\in\Lambda} \mathrm{r}(c)}{\mathrm{r}(\Lambda)}. \tag{7.1}$$

Remark that the combinatorial optimization problem in (7.1) is exponential in C. Even though real-life applications typically involve a small number of sensitive attributes (often only 2), applications involving intersectional fairness between different sensitive features may introduce a larger number of colors [BG18b; Kea+18; ML22b]. The PoF can be computed in time $\mathrm{poly}(C, |E|)$ whenever the underlying matroid possesses a **polynomial-time independence oracle** as is the case of transversal, graphic, and uniform matroids. Refer to Section 7.8 for the details.

In the following section we leverage Corollary 7.8 to tightly bound the PoF in various settings.

7.4 Bounding the Price of Fairness

7.4.1 Adversarial Price of Fairness

Our first result is to show that the price of opportunity fairness is always bounded, with a bound only depending linearly on C, the number of colors of the colored matroid, independent on the number of agents $|E|$ and the number of feasible allocations $|\mathcal{I}|$. Refer to Section 7.9.5 for the proof.

Theorem 7.9. *For any C-colored matroid M, we have $\mathrm{PoF}(M) \leqslant C - 1$, and this bound is tight.*

Theorem 7.9 implies the following remarkable result.

Corollary 7.10. *For any 2-colored matroid M, $\mathrm{PoF}(M) = 1$.*

Corollary 7.10 shows that whenever agents are divided in two groups, no social welfare loss is incurred due to the opportunity fairness constraint. For an alternative geometrical proof of Corollary 7.10, please refer to Section 7.9.6.

As an application example of Proposition 7.5, we can immediately obtain the following corollary.

Corollary 7.11. *Let $\gamma \in [0, 1]$. For any C-colored matroid M, we have $\mathrm{PoF}_\gamma(M) \leqslant \gamma \cdot (C - 1)$, and this bound is tight.*

Proof. This is immediate from Theorem 7.9, Proposition 7.5, and from remarking that the upper bound in Proposition 7.5 is increasing in $\mathrm{PoF}(M)$. □

Theorem 7.9 raises the question of whether tighter bounds can be obtained by restricting the resource allocation problem to specific subclasses of matroids. However, the bound is in fact tight for any class of matroids that contains either graphic or transversal matroids. Refer to Section 7.9.5 for the details.

7.4.2 Parametric Price of Fairness

The worst-case bound derived in the previous section relies on the existence of a specific polymatroid, as outlined in the proof of Theorem 7.9. Essentially, this requires an underlying structure where one group E_c can have a rank that grows arbitrarily large while the ranks of other groups remain bounded. This raises the question of whether more favorable guarantees for the price of opportunity fairness can be attained when all groups exhibit similar ranks. Proposition 7.12 provides an upper bound on PoF based on the ranks of the groups. The proof is detailed in Section 7.9.7.

Proposition 7.12. *For any C-colored matroid M, it holds,*

$$\mathrm{PoF}(M) \leqslant \frac{1}{2} \cdot \frac{\max_{c\in[C]} \mathrm{r}(c)}{\min_{c\in[C]} \mathrm{r}(c)} + \frac{C}{4} \cdot \left(\frac{\max_{c\in[C]} \mathrm{r}(c)}{\min_{c\in[C]} \mathrm{r}(c)}\right)^2 + \frac{1}{4C} \cdot \mathbb{1}\{C \text{ odd}\}.$$

Moreover, whenever all groups have the same rank, the resulting bound is tight.

The bound in Proposition 7.12 takes into account the shape of the polytope M. When all colors have the same rank, PoF scales as $C/4$. While this upper bound is smaller than the one stated in Theorem 7.9, the price of fairness remains linear with respect to the number of colors.

To complement the analysis, we consider another geometrical parameter related to the shape of M that can interpolate PoF between 1 and $C/4$. Intuitively, PoF is expected to be low when either there is no competition between groups, or the competition is extremely fierce and no group can be unilaterally allocated resources without damaging the allocation of others. Similar behavior has been observed for the Price of Anarchy in congestion games [Col+20], which approximates to 1 under both light and heavy traffic conditions. In the context of the price of opportunity fairness, the relevant problem complexity measure is the associated *independence index* that we define below.

Definition 7.13. We define the **independence index** of a polymatroid M as $\rho(M) := \frac{\mathrm{r}([C])}{\sum_{c\in[C]}\mathrm{r}(c)}$.

The independence index measures how close the maximal social welfare $\mathrm{r}([C])$ is to the social welfare of the *utopian allocation*, the allocation where each group E_c receives $\mathrm{r}(c)$ resources (which, in general, is not a feasible allocation). Note that ρ always falls within the interval $[1/C, 1]$, with $\rho = 1/C$ corresponding to complete competition between groups, and $\rho = 1$ corresponding to full independence between groups. These extreme values of the independence index impose a distinct shape on M, as illustrated in Figure 7.3.

Proposition 7.14. *Let M be a C-colored matroid. Suppose that for any c, c' in $[C]$, $\mathrm{r}(c) = \mathrm{r}(c')$. Whenever $\rho \in [1/C, 1/(C-1)]$, $\mathrm{PoF}(M) = 1$. Otherwise,*

$$\mathrm{PoF}(M) \leqslant \rho \max\left(\frac{C - \lfloor C\rho\rfloor + 1}{C\rho - \lfloor C\rho\rfloor + 1}, C - \lfloor C\rho\rfloor\right) \leqslant \rho((1-\rho)C + 1).$$

In addition, the first upper bound is jointly tight in ρ and C.

The proof of Proposition 7.14 is provided in Section 7.9.8, where Figure 7.11 illustrates a tight example of the first upper bound for a transversal matroid. Figure 7.2 illustrates both upper bounds from Proposition 7.14 for $C = 5$ groups with equal ranks. We observe that in both extremes, PoF tends towards 1. Notice that the second upper bound in Proposition 7.14, when maximized over ρ, aligns with the bound from Proposition 7.12 (when all groups have identical ranks). Therefore, the independence index interpolates the price of opportunity fairness between 1 and an order of $C/4$ when all groups have the same isolated social welfare.

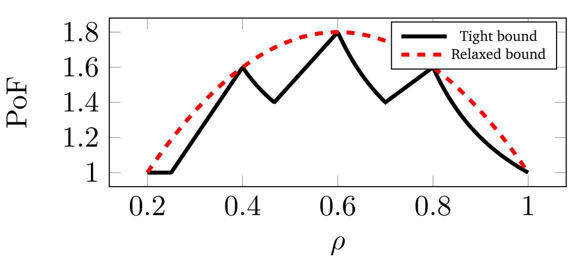


Fig. 7.2: PoF upper bounds stated in Proposition 7.14 for 5 groups with equal rank, with variable value of the independence index ρ.

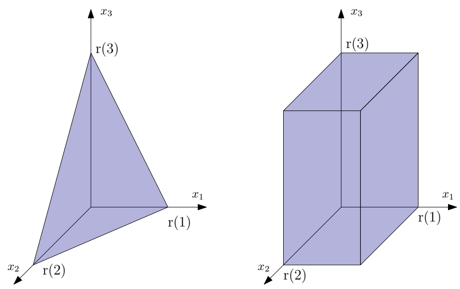


Fig. 7.3: Shape of M for extreme values of independence index. Left: $\rho = 1/C$, right: $\rho = 1$.

Worst-case analysis results stand out by their robustness. However, the particular matroid examples attaining the upper bounds (even under the extra structural assumptions) are rarely observed in real-life. Due to this, the following sections will be dedicated to analyzing PoF in random settings.

7.4.3 Semi-Random Price of Fairness (Random Coloring)

Our first random setting considers an adversarial matroid choice with a random group agents partition. Formally, we denote by Δ^C the simplex of dimension C, that is, the set of all vectors $p \in [0,1]^C$ such that $\sum_{c \in C} p_c = 1$. Given a vector $p \in \Delta^C$ without null entries, we create a random partition of a matroid $(E, \mathcal{I})$ (a coloring of the elements in E) by independently and identically assigning each element $e \in E$ to $c \in [C]$ with probability p_c. We denote by $M(p)$ the polymatroid obtained by the random coloring of the agents in E according to the vector p.

Let $(M_n(p))_{n \in \mathbb{N}}$ be a sequence of C-colored matroids over sets $(E^{(n)})$ such that $|E^{(n)}| = n$, randomly colored according to p, and $(\mathrm{r}_n)_{n \in \mathbb{N}}$ the associated sequence of rank functions. For each $c \in [C]$, suppose the following limit exists,

$$R(p_c) := \lim_{n \to \infty} \frac{\mathbb{E}_{p_c}[\mathrm{r}_n(c)]}{n}, \tag{7.2}$$

where $\mathbb{E}_{p_c}[\mathrm{r}_n(c)]/n$ represents the rescaled expected social welfare of group c. Remark that assuming convergence is not particularly restrictive as $\mathbb{E}_p[\mathrm{r}_n(c)]/n$ is bounded in $[0,1]$ and thus, it always admits a converging subsequence. We extend the previous definition to any subset $\Lambda \subseteq [C]$ by,

$$R\Big(\sum\nolimits_{c \in \Lambda} p_c\Big) := \lim_{n \to \infty} \frac{1}{n} \cdot \mathbb{E}_{(p_c)_{c \in \Lambda}}[\mathrm{r}_n(\cup_{c \in \Lambda} E_c)].$$

Finally, we will assume that $\mathrm{r}_n([C]) = \Omega(n)$, which ensures that the size of the optimal allocation grows with the size of the ground sets $E^{(n)}$. We first show that the price of opportunity fairness for large colored matroids is completely characterized by the function R.

Proposition 7.15. *If* $\liminf_{n \to \infty} \frac{r_n([C])}{n} > 0$*, then*

$$\mathrm{PoF}(M_n(p)) \xrightarrow[n \to \infty]{\mathrm{P}} \max_{\Lambda \subseteq [C]} \frac{R(1)}{\sum_{c \in [C]} R(p_c)} \cdot \frac{\sum_{c \in \Lambda} R(p_c)}{R(\sum_{c \in \Lambda} p_c)}, \tag{7.3}$$

where $\xrightarrow{\mathrm{P}}$ *denotes convergence in probability. Moreover, the function* R *is such that* $R(0) = 0$*,* R *is concave, non-decreasing, and* 1*-Lipschitz. Conversely, any function with* *these three properties can be realized as the limit of a sequence of limit-matroid derived functions R'_m, i.e.* $\|R - R'_m\|_\infty \xrightarrow[m\to\infty]{} 0$.

Proof sketch. We first show that, as R is the multi-linear extension of a submodular function, it satisfies the aforementioned properties. Using the concavity of R, we show that whenever $\mathrm{r}([C])$ is large, with high probability, $\mathrm{r}(c)$ is large as well. Then, by using McDiarmid's concentration inequality, the convergence in probability is concluded. The approximation result is proved by constructing a family of simple functions from specific sequences $(M_n)_n$ whose closed convex hull is equal to the desired set of functions. The full proof is included in Section 7.9.9. $\square$

The first property of Proposition 7.15 shows that upper bounding the right-hand side of Equation (7.3) yields an upper bound on PoF. The second part shows an equivalence between sequences of C-colored matroids and the set of concave, non-decreasing, 1-Lipschitz functions, which we denote by $\mathcal{R}$. Therefore, we can shift the problem of bounding the price of opportunity fairness of C-colored matroids to bounding the right-hand side of Equation (7.3) over all functions in $\mathcal{R}$. We aim to find a bound that depends only on $\max_{c\in[C]} p_c$. The following theorem provides the exact solution:

Theorem 7.16. *Fix $\pi \in [1/C, 1]$ and consider $\Delta^C_\pi := \{p \in \Delta^C \mid \max_{c\in[C]} p_c = \pi\}$ the set of probability distributions with maximum value of π. It follows that*

$$\max_{p\in\Delta^C_\pi} \max_{R\in\mathcal{R}} \max_{\Lambda\subseteq[C]} \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)} = \max_{\lambda\in[C]} \psi_\lambda\left(\frac{1-(C-\lambda)\pi}{C}\right) \leqslant \min(C - \frac{1}{\pi}, 1), \tag{7.4}$$

where $\psi_\lambda : [-\lambda, \frac{1}{C}] \to \mathbb{R}$, for each $\lambda \in [C]$, is given by

$$\psi_\lambda(q) = \begin{cases} \lambda & \textit{if } q \in [-\lambda, 0]\,, \\ \frac{\lambda}{(\lambda C q-1)^2} \cdot \left(1+q(1-2\lambda)+C(\lambda-2+\lambda q)q-2\sqrt{(\lambda-1)(C-1)(1-Cq)(1-\lambda q)q}\right) & \textit{if } q \in \left(0, \frac{(\lambda-1)}{\lambda(C-1)}\right]\,, \\ 1 & \textit{if } q \in \left(\frac{(\lambda-1)}{\lambda(C-1)}, \frac{1}{C}\right]\,. \end{cases}$$

Proof sketch. The left-hand side of Equation (7.4) corresponds to an infinite-dimensional combinatorial optimization problem, for which classical techniques are difficult to apply. We address this by designing transformations that map a generic instance (p, Λ, R) to a new instance (p', Λ', R') with a *larger Price of Fairness*. These transformations include linearizing R over subintervals of $[0, 1]$, using Karamata's inequality to average the probabilities of the indices in Λ, modifying Λ to ensure that the maximum-probability coordinate $c^* \in [C]$ of p lies outside Λ, among others. Iteratively applying these transformations, we manage to reduce the original triple optimization problem, at a fixed size $\lambda = |\Lambda|$, to a *single-variable* optimization problem, which can then be solved via first-order conditions to obtain ψ_λ, the worst-case value for that size. We then take the maximum over all possible values of λ. The complete proof is provided in Section 7.9.10. □

We have reduced a complex optimization problem to a simple closed-form formula with a nice interpretation, the ψ_λ corresponding to the worst-case price of opportunity fairness for a fixed size $\lambda = |\Lambda|$. It may intuitively seem that the Price of Fairness should always be 1, as randomly coloring the ground set is equivalent to first drawing a random number of agents per group and then placing them uniformly on the ground set, seemingly ensuring group-independent opportunities. Nonetheless, a striking conclusion from Theorem 7.16 is that **the price of opportunity fairness can still exceed 1, even when all agents are treated identically**. In particular, whenever $\max_{c\in[C]} p_c$ is larger than $1/2$, the worst-case PoF is at least $C-2$, as observed in Figure 7.4.

On the other hand, Theorem 7.16 provides meaningful PoF bounds whenever $\max_{c\in[C]} p_c \leqslant 1/2$. More importantly, Theorem 7.16 immediately implies the following corollary.

Corollary 7.17. *Whenever* $\max_{c\in[C]} p_c \leqslant 1/(C-1)$*,* $\mathrm{PoF}(M_n)$ *converges in probability towards 1.*

In other words, for matroid allocation problems in large markets, there is no loss of social welfare due to opportunity fairness as long as no group is overrepresented. This is quite striking as this may contradict the intuition that unfairness stems from the presence of small protected groups that must be catered to, sacrificing the welfare of larger groups. The above corollary shows that even with the presence of an arbitrarily small group, there might be no social welfare loss when being fair. Instead, it is the presence of a single overwhelming group which makes resources hard to fairly allocate, for the specific notion of opportunity fairness.

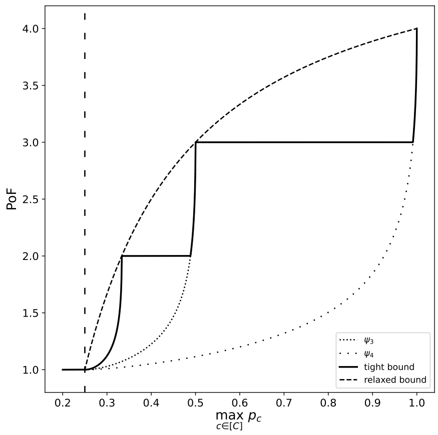


Fig. 7.4: Theorem 7.16 for $C = 5$ groups: worst-case PoF, ψ_λ for $\lambda \in \{3, 4\}$, and the relaxed bound $C - 1/\pi$.

We have shown how to bound the price of opportunity fairness in the semi-random setting by reducing the combinatorial optimization problem to make it tractable. However, considering the underlying matroid $(E, \mathcal{I})$ to be fixed may still be a pessimistic assumption for some real-world applications. For this reason, we study next a setting where both the matroid and the colors are drawn randomly.

7.4.4 Random Graphs Price of Fairness

A first possibility to construct random matroids is to uniformly pick a matroid among the $2^{2^{n-O(\log(n))}}$ possible matroids for a ground set of size n, but this would mix different resource allocation problems. Instead, we will focus on sub-classes of matroids, where specific distributions over the matroids are already well established: we analyze random graphs, in particular Erdös-Rényi random graphs. We consider graphic matroids (finding the largest forest in a graph), and transversal matroids (finding the largest matching in a bipartite graph)—see Section 7.6 for the formal definitions.

Given $n \in \mathbb{N}$ and $q \in [0,1]$, we consider the Erdös-Rényi random graph $\mathrm{G}_{n,q} := ([n], A)$ with n nodes such that for any $i, j \in [n]$, the edge (i, j) belongs to A independently with probability q. Given a random graph $\mathrm{G}_{n,q} = ([n], A)$ and $p \in \Delta^C$, we consider a p **randomly colored random graphic matroid**, denoted $\mathrm{G}_{n,q}(p)$. Remark that the random coloring process and the random edges connections are done independently from one another.

Proposition 7.18. *Let $\omega = \omega(n)$ be a function such that $\omega(n) \to \infty$. Whenever $q \leqslant 1/(\omega n)$ or $q \geqslant \omega/n$, for any $p \in \Delta^C$, $\mathrm{PoF}(\mathrm{G}_{n,q}(p))$ converges to 1 with high probability as n grows.*

Given $n \in \mathbb{N}$, $\beta \in (0,1)$ such that $\beta n \in \mathbb{N}$, and $q \in [0,1]$, we consider the random bipartite graph $\mathrm{B}_{n,\beta,q} := ([n], [\beta n], A)$, where for any $i \in [n], j \in [\beta n]$, the edge $(i, j) \in A$ independently with probability q. Given a random Erdös-Rényi bipartite graph $\mathrm{B}_{n,\beta,q}$ and $p \in \Delta^C$, we consider a p **randomly colored random transversal matroid** denoted $\mathrm{B}_{n,\beta,q}(p)$. Recall, the coloring process and the edges are drawn independently between them.

Proposition 7.19. *Let $\omega = \omega(n)$ be a function such that $\omega(n) \to \infty$ arbitrarily slow as $n \to \infty$. Whenever $q \leqslant 1/(\omega n^{3/2})$ or $q \geqslant \omega \log(n)/n$, for any $p \in \Delta^C$, $\mathrm{PoF}(\mathrm{B}_{n,\beta,q}(p))$ converges to 1 with high probability as n grows.*

Conclusions and Limitations

We address matroid allocation problems under a novel group-fairness notion—opportunity fairness—, and prove tight bounds for the Price of Fairness, i.e., the loss of social welfare due to the fairness restrictions, in multiple settings from adversarial to fully random. Our model has two main limitations. First, integral allocations are only well approximated by fractional ones in large markets, which corresponds well to our motivating examples (e.g., job market) but may not always hold. Second, we considered the allocation cardinality as our notion of social welfare, i.e., each allocation has the same weight. The extension to weighted matching is straightforward if weights are assigned at the level of groups as it simply skews the polymatroid. However, individual-level weights would break the anonymity property and the polymatroid nature of M, hence new techniques would be required.

7.5 Further Related Works

Fairness notions: The study of fair algorithmic decision-making cover a broad range of fields, with fair division [Ste49] in economics that focuses on concepts such as envy-freeness [Car+19; Lip+04; Var74; Wel85] and maximin fairness [Sen17], and machine learning that emphasizes statistical fairness notions like group fairness [Con+19; Fre+23; S S+21], including demographic parity [LC23]. Our new notion of opportunity fairness is inspired from both Equality of Opportunity [Har+16c] and Kalai-Smorodinsky fairness [KS75; NPP17]. Equality of Opportunity aims for *true-positive* rates to be independent of sensitive attributes; for opportunity fairness, this translates to ensuring that the resources allocated to a group are proportional to its opportunity level, the maximum allocation it could receive if it were considered in isolation. On the other hand, Kalai-Smorodinsky fairness requires maximizing the ratio $|S \cap E_c| / \max_{S' \in \mathcal{I}} |S' \cap E_c|$. While any maximum-size opportunity fair allocation satisfies Kalai-Smorodinsky fairness, the reverse does not necessarily hold, making our fairness notion more restrictive.

The Price of Fairness: The Price of Fairness was concurrently introduced by [BFT11], who focus on maximin fairness and proportional fairness, and [Car+12] who prioritize equitability and envy-freeness. They provided bounds for each fairness notion depending on the number of agents. Subsequently, [NPP17] studied the price of fairness under the Kalai-Smorondinsky fairness notion for the *subset sum problem* and [DPS14] studied the price of fairness in kidney-exchange. The concept of price of fairness has been extended to others research domains such as supervised machine learning [Haa19; GLR20; MW21] where the cost of fairness is studied on different prediction tasks. In this article, we initiate the study of the Price of Fairness under opportunity fairness. Unlike equitability, which requires identical allocations across groups, opportunity fairness is a more robust notion in the context of price of fairness. When some groups are inherently unable to receive resources, enforcing equitability leads to significant welfare loss, whereas we show that the price of fairness under opportunity fairness remains bounded. More importantly, we go beyond the traditional adversarial worst-case analysis, and instead consider more structured inputs, in the vein of [Rou20], allowing for an average case analysis that better reflect trade-offs in real-world instances.

Fair Matroid Allocation Problems: The main objective of the matroid and fairness literature, initiated by [Chi+19], is to efficiently approximate maximum size fair allocations. Subsequent works extend this framework to submodular function optimization under fairness and matroid constraints [El +20; El +24; TY23; YT23], while [Ban+23] study the computational complexity of finding optimal proportionally fair matching for more than two groups. We remark that maximum size opportunity fair fractional allocations can be computed efficiently whenever the underlying matroid possesses a polynomial-time independence oracle ([Sch+03], Chapter 44), as is the case of bipartite matching and communication network problems.

Fair matching: Recent work in matching have increasingly examined the impact of different fairness notions on matching mechanisms and how to design fair algorithms. [CLP22; Dev+23] and [KK23] examine the relation between fairness and stability in matching. Additionally, fairness in online matching has been studied in various contexts, including waiting time, equality of opportunity, and fairness constraints on the offline side of the market [Chu16; Esm+22; Hos+23; MXX21; S S+21]. Our work contributes to this growing literature by introducing a new fairness notion, opportunity fairness, that captures the structural constraints of allocation problems.

7.6 Matroid classes

The three examples introduced in Section 7.1 can be modeled as matroid allocation problems. Indeed,

1. The bipartite matching problem corresponds to a **transversal matroid**: Let $G = (U, V, A)$ be a bipartite graph. For a matching $\mu \subseteq A$ we denote $\mu(U) := \{u \in$

$U \mid \exists v \in V, (u,v) \in \mu\}$. The pair $(U, \mathcal{I})$, with $\mathcal{I} := \{\mu(U), \forall \mu \subseteq A$ matching$\}$, is called a transversal matroid.

2. The communication network problem corresponds to a **graphic matroid**: Let $G = (U, A)$ be a graph. The pair $(A, \mathcal{I})$, with $\mathcal{I} := \{S \subseteq A \mid S$ is acyclic$\}$, is called a graphic matroid.

3. The constitutional commission problem corresponds to a **uniform matroid**. Let E be a finite set and $b \in \mathbb{N}$. The b-uniform matroid is the pair $(E, \mathcal{I})$, such that, $\mathcal{I} := \{S \subseteq E \mid |S| \leqslant b\}$.

7.7 Null integral opportunity fair allocations

Figure 7.5 illustrates an example on a graphic matroid[3] for two groups. The figure on the right shows the integer feasible allocations (the blue dots) and the set of opportunity fair allocations (the orange line), whose only intersection is at the origin.

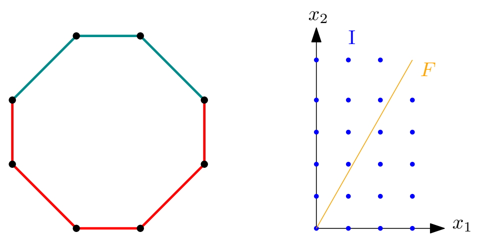


Fig. 7.5: Graphic matroid example showing that integrality can lead to null opportunity fair allocations. (Left) Graph defining a colored graphic matroid with two groups. Group 1 is denoted by the green edges while Group 2 by the red edges. It follows that $r(1) = 5$, $r(2) = 3$, and $r(\{1,2\}) = 7$. (Right) Set of integer feasible allocations (blue dots) and set of opportunity fair allocations (orange line), whose only intersection is at the origin.

7.8 Efficient computation of Opportunity fair allocations of maximum size

Whenever a matroid $(E, \mathcal{I})$ possesses a poly-time independent oracle, that is, an oracle that for any subset S of E tells if S is an independent set or not in polynomial

[3]Allocations within a graphic matroid correspond to acyclic subgraphs of the given graph.

time, we can construct a poly-time separation oracle for the corresponding polymatroid in $\mathbb{R}^{|E|}$ [Sch+03][Theorem 40.4]. A separation oracle, given a vector $x \in \mathbb{R}^{|E|}$ and the polymatroid

$$Q := \left\{ x \in \mathbb{R}_+^{|E|} \mid \sum_{e \in S} x_e \leqslant \mathrm{rank}(S), \forall S \subseteq E \right\},$$

either tells if x belongs to Q or outputs a separating hyperplane. Starting from the poly-time separation oracle of a matroid, we can obtain a poly-time separation oracle for set of feasible fractional allocation of the corresponding colored matroid (or polymatroid). Therefore, we can compute an opportunity fair allocation of maximum size by solving the following problem,

$$\begin{aligned}
\max \quad & \sum_{c \in [C]} x_c \\
\text{s.t. } x_c = & \sum_{e \in E : e \in c} y_e && \forall c \in [C] \\
& \frac{x_c}{\mathrm{rank}(c)} = \frac{x_{c'}}{\mathrm{rank}(c')}, && \forall c, c' \in [C] \\
& (y_e)_{e \in E} \text{ is a feasible fractional allocation} \\
& x \in [0,1]^C, y \in [0,1]^{|E|}.
\end{aligned}$$

Indeed, the previous problem can be solved in polynomial time using the ellipsoid method provided with the poly-time separation oracle of the polymatroid and the (trivial) poly-time separation oracle of the polytope of fairness constraints.

7.9 Missing proofs

7.9.1 Proof of Proposition 7.4

Proposition 7.4. For any fractional opportunity fair maximum size allocation $x \in F$, there always exists a random allocation X, such that, $\mathbb{E}[X] = x$ and for all realizations, X is feasible, integral, $\left(1 - \frac{2C}{\min_c r(c)}\right)$-opportunity fair, and $\|X - x\|_1 \leqslant C$.

Proof. Let M be a set of fractional allocations and x be an optimal fair allocation. Consider the translated unit cube $K = \lfloor x \rfloor + [0,1]^C$ where $\lfloor x \rfloor = (\lfloor x_1 \rfloor, \ldots, \lfloor x_C \rfloor)$, which contains x and all the integer vertices closest to x. We aim to prove that $K \cap M$ yields integer extreme points.

Let $M' := \mathbb{R}_+^C \cap (M - \lfloor x \rfloor)$ to be the positive quadrant of M translated by $\lfloor x \rfloor$. It holds

$$
\begin{aligned}
M' &= \{y \in \mathbb{R}_+^C \mid \sum_{c\in\Lambda} y_c + \lfloor x_c \rfloor \leqslant r(\Lambda), \forall \Lambda \subset [C]\} \\
&= \{y \in \mathbb{R}_+^C \mid \sum_{c\in\Lambda} y_c \leqslant r(\Lambda) - \sum_{c\in\Lambda} \lfloor x_c \rfloor, \forall \Lambda \subset [C]\}.
\end{aligned}
$$

Consider the function $f(\Lambda) := r(\Lambda) - \sum_{c\in\Lambda} \lfloor x_c \rfloor$. Since $x \in M$, so is $\lfloor x \rfloor$, hence $f(\Lambda) \geqslant 0$, and f takes only integer values as both r and $\lfloor x \rfloor$ take integer values. Moreover, f is submodular as both r and the application $\Lambda \mapsto \sum_{c\in\Lambda} x_c$ are submodular. Therefore, M' is an integral polymatroid.

The unit cube being also an integral polymatroid, by Theorem 35 of [Edm70], $[0,1]^C \cap M'$ has integral vertices, and thus so it does,

$$
\lfloor x \rfloor + ([0,1]^C \cap M') = K \cap M.
$$

It follows that, since $x \in K \cap M$, which is a convex set (as both K and M are convex sets), x can be described as the convex combination of the extreme points of $K \cap M$. Since all extreme points of $K \cap M$ are integral, they are, in particular, a subset of the extreme points of K. By Carathéodory's Theorem, there exists $X^{(1)}, \ldots, X^{(k)}$ vertices of $K \cap M$ for $k \leqslant C+1$, and $(p_1, \ldots, p_k) \in [0,1]^{C+1}$ with $\sum_{i=1}^k p_i = 1$, such that $x = \sum_{i=1}^k p_i X^{(i)}$.

The randomized rounding X then is defined by drawing $X^{(i)}$ with probability p_i. Note that for any realization of $X^{(i)}$ of X, $X^{(i)}$ is feasible and integral by construction. Moreover,

$$
\mathbb{E}[X] = \sum_{i=1}^k p_i X^{(i)} = x,
$$

and since $X^{(i)}$ for any $i \in \{1, ..., k\}$, and x belong to $[0,1]^C$, it always follows that $\|X^{(i)} - x\|_1 \leqslant C$. Regarding the γ-opportunity fairness, recall that x is 1-opportunity fair, so there exists $\alpha \in [0,1]$ such that $x = \alpha(r(1), \ldots, r(C))$. Moreover, because x is optimal, $\alpha \geqslant 1/C$. Indeed, either $\mathrm{PoF} = 1 = r([C])/(\alpha \sum_c r(c))$ which yields $\alpha = r([C])/\sum_c r(c) \geqslant r([C])/(Cr([C])) = 1/C$, or $\mathrm{PoF} > 1$ in which case $(r(1), \ldots, r(C))/(C-1)$ is feasible (see proof of Theorem 7.9) which implies, by definition, that $\alpha \geqslant (C-1)$. For any $i, j \in [C]$, it follows,

$$
\frac{X_i/r(i)}{X_j/r(j)} \geqslant \frac{(x_i - 1)/r(i)}{(x_j + 1)/r(j)} \geqslant \frac{\alpha - 1/\min_c r(c)}{\alpha + 1/\min_c r(c)} = 1 - \frac{2}{\alpha \min_c r(c) + 1} \geqslant 1 - \frac{2C}{\min_c r(c)},
$$

concluding the proof. $\square$

7.9.2 Proof of Proposition 7.5

Proposition 7.5. Let M be a C-colored matroid. Then, either $\gamma < \max_{x\in \mathrm{F}} \min_{c\in[C]} x_c/r_c$ and $\mathrm{PoF}_\gamma(M) = 1$, or

$$\gamma \cdot \mathrm{PoF}(M) \leqslant \mathrm{PoF}_\gamma(M) \leqslant \frac{\gamma \cdot \mathrm{PoF}(M)}{1-(1-\gamma)^{\frac{((C-1)-\mathrm{PoF}(M))}{C-2}}}.$$

Proof. **Lower bound.** Let $x = \arg\max_{y\in F} \|y\|_1 = \alpha(r(1), ..., r(C))$, where $\alpha \in [0,1]$. If x is maximal, $\mathrm{PoF}_\gamma(M) = \mathrm{PoF}(M) = 1$. Otherwise, there exists $\Lambda \subsetneq [C]$ such that $x \in \arg\max_{y\in M} \sum_{c\in\Lambda} y_c$. Define x' as $x'_i = x_i$ for all $i \in \Lambda$ and $x'_i = \frac{1}{\gamma} x_i$ for all $i \notin \Lambda$.

Let us denote by F_γ the γ-opportunity fair feasible set. First, we prove that $\|x'\|_1 \geqslant \max_{F_\gamma} \|x\|_1$. Suppose there exists $x'' \in F_\gamma$ a γ-opportunity realizable point such that $\|x''\|_1 > \|x'\|_1$. Let $\Gamma = \{i \in [C], x''_i > x'_i\}$. Since

$$\sum_\Lambda x''_i \leqslant r(\Lambda) = \sum_\Lambda x'_i,$$

it follows that, $\sum_{\Lambda^c} x''_i > \sum_{\Lambda^c} x'_i$, and thus $\Gamma \cap \Lambda^c \neq \emptyset$. Let $i \in \Gamma \cap \Lambda^c$, then $x''_i > x'_i = \frac{1}{\gamma} x_i$. Moreover, there must exist $j \in \Lambda$ such that $x''_j \leqslant x_j$ since $x \in \arg\max_M \sum_{c\in\Lambda} y_c$. Therefore,

$$\frac{x''_i}{r(i)} > \frac{1}{\gamma} \cdot \frac{x_i}{r(i)} = \frac{1}{\gamma} \cdot \frac{x_j}{r(j)} \geqslant \frac{1}{\gamma} \cdot \frac{x''_j}{r(j)},$$

which contradicts x'' being γ-opportunity fair. It follows,

$$\|x'\|_1 = \sum_{i\in\Lambda} x_i + \frac{1}{\gamma} \sum_{i\in\Lambda^c} x_i = \alpha\Big(\sum_{i\in\Lambda} r(i) + \frac{1}{\gamma} \sum_{i\in\Lambda^c} r(i)\Big),$$

and therefore,

$$\frac{\mathrm{PoF}(M)}{\mathrm{PoF}_\gamma(M)} = \frac{\max_{y\in F_\gamma} \|y\|_1}{\max_{y\in F} \|y\|_1} \leqslant \frac{\|x'\|_1}{\|x\|_1} = \frac{\sum_{i\in\Lambda} r(i) + \frac{1}{\gamma}\sum_{i\in\Lambda^c} r(i)}{\sum_{i\in[C]} r(i)} \leqslant \frac{1}{\gamma},$$

thus, $\mathrm{PoF}_\gamma(M) \geqslant \gamma \mathrm{PoF}(M)$.

Upper bound. As before, let $x = \arg\max_{y\in F} \|y\|_1 = \alpha(r(1), ..., r(C))$. We suppose that x is not optimal, otherwise $\mathrm{PoF}_\gamma = \mathrm{PoF} = 1$. Let x' be a point in the Pareto frontier that Pareto dominates x. By the polymatroid characterization, it holds that x is minimal, i.e. $\|x\|_1 = r([C])$. For $\lambda \in \mathbb{R}_+$, consider the combination $z = \lambda x' + (1-\lambda)x$. Because x' Pareto dominates x, we have $x'_i \geqslant x_i$ for all $i \in [C]$, hence $z_i \geqslant x_i$. By definition of $r(i)$ we also have that $x'_i \leqslant r(i)$, thus $z_i \leqslant \lambda r(i) + (1-\lambda)x_i$. This implies that

$$\frac{z_i/r(i)}{z_j/r(j)} \geqslant \frac{x_i/r(i)}{\lambda + (1-\lambda)(x_j/r(j))} = \frac{\alpha}{\lambda + (1-\lambda)\alpha}.$$

In particular, z is γ-opportunity fair whenever $\alpha/(\lambda + (1-\lambda)\alpha) \geqslant \gamma$, i.e., whenever $\lambda \leqslant \frac{\alpha(1-\gamma)}{\gamma(1-\alpha)}$. If $\alpha(1-\gamma)/(\gamma(1-\alpha)) > 1$, which is equivalent to $\gamma < \alpha$, then taking $\lambda = 1$ yields that x' is γ-fair, and it is already optimal so $\mathrm{PoF}_\gamma = 1$. In the following we let $\gamma \geqslant \alpha$. Let us take $\lambda = \alpha(1-\gamma)/(\gamma(1-\alpha)) \leqslant 1$ such that z in M by convex combination. Because z is γ-opportunity fair and feasible, $z \in F_\gamma$. Hence

$$\begin{aligned}
\mathrm{PoF}_\gamma(M) = \frac{r([C])}{\max_{y \in F_\gamma} \|y\|_1} \leqslant \frac{r([C])}{\|z\|_1} &= \frac{r([C])}{\lambda \sum_{c\in[C]} x'_c + (1-\lambda)\sum_{c\in[C]} x_c} \\
&= \frac{\mathrm{PoF}(M)}{\lambda \mathrm{PoF}(M) + (1-\lambda)\cdot 1} \\
&= \frac{\mathrm{PoF}(M)}{\frac{\alpha(1-\gamma)}{\gamma(1-\alpha)}\mathrm{PoF}(M) + \frac{\gamma(1-\alpha)-\alpha(1-\gamma)}{\gamma(1-\alpha)}} \\
&= \frac{\gamma(1-\alpha)\mathrm{PoF}(M)}{\alpha(1-\gamma)\mathrm{PoF}(M) + \gamma - \alpha} \\
&\leqslant \frac{\gamma(1-\frac{1}{C-1})\mathrm{PoF}(M)}{\frac{1}{C-1}(1-\gamma)\mathrm{PoF}(M) + \gamma - \frac{1}{C-1}} \\
&= \frac{\gamma \mathrm{PoF}(M)}{1-(1-\gamma)\frac{((C-1)-\mathrm{PoF}(M))}{C-2}},
\end{aligned}$$

where we used that $\alpha \geqslant 1/(C-1)$ as otherwise $\mathrm{PoF}(M) = 1$. □

7.9.3 Proof of Proposition 7.7

Proposition 7.7. Let $((E_c)_{c\in[C]}, \mathcal{I})$ be a colored matroid with rank function r and set of feasible fractional allocations M. Then, M is the polymatroid associated to the function r, i.e.,

$$M = \{x \in \mathbb{R}_+^C \mid \textstyle\sum_{c\in\Lambda} x_c \leqslant \mathrm{r}(\Lambda), \forall \Lambda \subseteq [C]\}.$$

Proof. Let $((E_c)_{c\in[C]}, \mathcal{I})$ be a C-colored matroid. It is sufficient to show that

$$\mathrm{I} := \{(|S \cap E_1|, \ldots, |S \cap E_C|) \in \mathbb{N}^C \mid S \in \mathcal{I}\},$$

the set of integer feasible allocation, is a discrete polymatroid, and to conclude by taking the convex hull, as an integral polymatroid is equivalent to the convex hull of a discrete polymatroid [Edm70]. We prove this by showing that I satisfies equivalent conditions for a set to be a discrete polymatroid [HH11], which are:

1. For any $x \in \mathbb{N}^C$ and $y \in \mathrm{I}$ such that $x \leqslant y$ (component wise), $x \in \mathrm{I}$.
2. For any $x, y \in \mathrm{I}$, with $\|x\|_1 < \|y\|_1$, there exists $c \in [C]$ such that $x_c < y_c$ and $x + \vec{e}_c \in \mathrm{I}$, where $\vec{e}_c$ is the canonical vector with value 1 at the c-th entry and 0 otherwise.

The first property is a direct consequence of the hereditary property of $\mathcal{I}$. Let $x, y \in \mathrm{I}$ such that $\|x\|_1 < \|y\|_1$. Let $A_x, A_y \in \mathcal{I}$ be two independent sets such that $(|A_x \cap E_1|, \ldots, |A_x \cap E_C|) = x$ and $(|A_y \cap E_1|, \ldots, |A_y \cap E_C|) = y$. In particular, $|A_x| < |A_y|$. By the augmentation property, there exists $e \in A_y \setminus A_x$ such that $A_x \cup \{e\} \in \mathcal{I}$. Let c be the group of e. If $x_c < y_c$, the proof is over. Suppose, otherwise, that $x_c \geqslant y_c$. Let $e' \in A_x$ such that $e' \in E_c$ as well, and define $A'_x = A_x \cup \{e\} \setminus \{e'\}$. By the hereditary property, $A'_x \in \mathcal{I}$ and, by construction, $(|A'_x \cap E_1|, \ldots, |A'_x \cap E_C|) = x$. Applying again the augmentation property, there exists $e'' \in A_y \setminus A'_x$ such that $A'_x \cup \{e''\} \in \mathcal{I}$. Notice that $e'' \notin \{e, e'\}$. The proof is concluded by repeating the same argument until finding an element in some E_c such that $x_c < y_c$. The procedure stops after a finite amount of iteration as at every iteration the element in A_y obtained from the augmentation property must be different to all previous ones as well to all replaced elements in A_x. □

7.9.4 Proof of Corollary 7.8

Corollary 7.8. The price of opportunity fairness of a polymatroid M is given by,

$$\mathrm{PoF}(M) = \frac{\mathrm{r}([C])}{\sum_{c \in [C]} \mathrm{r}(c)} \cdot \max_{\Lambda \subseteq [C]} \frac{\sum_{c \in \Lambda} \mathrm{r}(c)}{\mathrm{r}(\Lambda)}.$$

Proof. Let x^* be a maximum size opportunity fair allocation. The opportunity fair requirement implies that x^* belongs to the line $t \cdot (\mathrm{r}(1), \ldots, \mathrm{r}(C))$ for $t > 0$. Let $t^* > 0$ such that $x^* = t^* \cdot (\mathrm{r}(1), \ldots, \mathrm{r}(C))$. Since x^* is a feasible fractional allocation, Proposition 7.7 implies that for any $\Lambda \subseteq [C]$,

$$t^* \sum_{c \in \Lambda} \mathrm{r}(c) \leqslant \mathrm{r}(\Lambda).$$

It follows that

$$t^* = \min_{\Lambda \subseteq [C]} \frac{\mathrm{r}(\Lambda)}{\sum_{c\in\Lambda} \mathrm{r}(c)},$$

and, in particular, that,

$$\mathrm{PoF}(M) = \frac{\mathrm{r}([C])}{t^* \sum_{c\in[C]} r(c)} = \frac{\mathrm{r}([C])}{\sum_{c\in[C]} \mathrm{r}(c)} \cdot \max_{\Lambda\subseteq[C]} \frac{\sum_{c\in\Lambda} \mathrm{r}(c)}{\mathrm{r}(\Lambda)},$$

concluding the proof. □

7.9.5 Proof of Theorem 7.9

Theorem 7.9. For any C-colored matroid M, we have $\mathrm{PoF}(M) \leqslant C-1$, and this bound is tight.

Proof. Let M be a C-dimensional polymatroid. Let Λ^* be the maximizer of Equation (7.1). Whenever $\Lambda^* = [C]$ it holds $PoF(M) = 1 \leqslant C-1$. Suppose then that $\Lambda^* \subsetneq [C]$. It follows,

$$\mathrm{PoF}(M) = \frac{\mathrm{r}([C])}{\mathrm{r}(\Lambda^*)} \cdot \frac{\sum_{c\in\Lambda^*} \mathrm{r}(c)}{\sum_{c\in[C]} \mathrm{r}(c)}.$$

Since $\mathrm{r}(\Lambda^*) \geqslant \max_{c\in\Lambda^*} \mathrm{r}(c) \geqslant \frac{1}{|\Lambda^*|} \sum_{c\in\Lambda^*} \mathrm{r}(c)$, it follows that,

$$\mathrm{PoF}(M) \leqslant |\Lambda^*| \cdot \frac{\mathrm{r}([C])}{\sum_{c\in\Lambda^*} \mathrm{r}(c)} \cdot \frac{\sum_{c\in\Lambda^*} \mathrm{r}(c)}{\sum_{c\in[C]} \mathrm{r}(c)} \leqslant |\Lambda^*| \leqslant C-1,$$

where we have used that $\mathrm{r}([C]) \leqslant \sum_{c\in[C]} \mathrm{r}([C])$ because r is submodular and thus sub-additive, and $|\Lambda^*| \leqslant C-1$ as $\Lambda^* \neq [C]$.

To show that this bound is tight, we exhibit a sequence of C-dimensional polymatroids for which the bound is tight in the limit. Consider a bipartite graph as in Figure 7.6. Let E_1 be independently connected to r_1 nodes, and $E_2, \ldots, E_C$ be completely connected to the same r_2 nodes. The last $C-1$ groups are in competition for resources while the first group suffers no competition. It holds $\mathrm{r}(1) = r_1$, while for any $\Lambda \subseteq [C] \setminus \{1\}$, $\mathrm{r}(\Lambda) = r_2$, and $\mathrm{r}(\Lambda \cup \{1\}) = r_1 + r_2$. In particular, it follows that Λ^*, the maximizer of Equation (7.1), is given by $\Lambda^* = [C] \setminus \{1\}$, and,

$$\mathrm{PoF}(M) = \frac{r_1 + r_2}{r_1 + (C-1)r_2} \cdot \frac{(C-1)r_2}{r_2} = \frac{r_1 + r_2}{r_1 + (C-1)r_2} \cdot (C-1) \xrightarrow[r_1\to\infty]{} C-1.$$ □

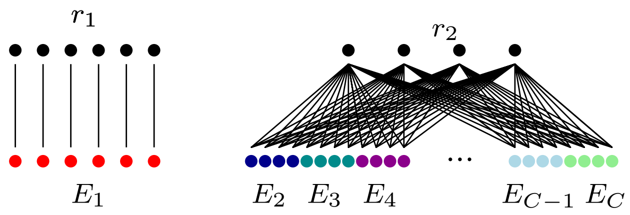


Fig. 7.6: Transversal polymatroid that makes the PoF bound of $C - 1$ tight (Theorem 7.9). Group 1 is totally independent of the rest of the groups, while groups $\{2, 3, ..., C\}$ compete for the same resources.

The proof of Theorem 7.9 shows that for transversal matroid, i.e. bipartite matching, our main application, the bound is tight. Regarding **graphic matroids**, the Walecki construction [Als08] which states that any clique of $2C - 1$ vertices has a decomposition into $C - 1$ disjoint Hamiltonian cycles, allows to design a tight example for the PoF upper bound. Indeed, associate each Hamiltonian cycle to a color $c \in [C - 1]$ and add one extra group of edges with color C as illustrated in Figure 7.7. It is not hard to see that this construction achieves a PoF equal to $C - 1$.

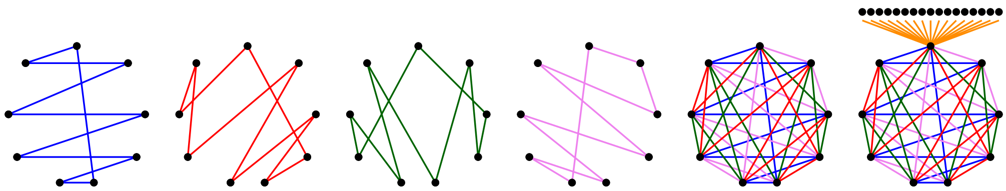


Fig. 7.7: Graphic polymatroid tight example for worst-case PoF equal to $C-1$ (Theorem 7.9), based on the Walecki construction, for 5 groups. Groups are represented by colored edges. The first 4 figures show the edges of each group, while the clique corresponds to their union. The final figure considers a 5-th group that is totally independent on the rest, leading to a similar construction as in Figure 7.6.

The same transversal matroid example can be used for **partition matroid**. This also implies the tightness of this bound for larger sub-classes of matroid which include either graphic or partition matroids, such as **linear** or **laminar matroids**. For **uniform matroids**, in exchange, it is immediate to see that $\mathrm{PoF}(M) = 1$: if the opportunity fairness constraint is violated, it is always possible to take the excessive (potentially fractional) resources from over-represented colors and give them to the under-represented ones. It remains open to prove if intermediate cases (not 1 nor $C - 1$) exist for some family of matroids.

7.9.6 Alternative proof of Corollary 7.10

Corollary 7.10 states that no loss of social welfare is incurred by imposing opportunity fairness when agents are divided into two groups. The same conclusion can be easily proven from a geometric point of view, as the line directed by (r_1, r_2) necessarily intersects with the Pareto frontier, which corresponds to the set of social welfare maximizing allocations and therefore, it is directed by $(1, -1)$. For $C > 2$, the property does not hold anymore, as proven by the tight $C-1$ bound, since the line directed by $(r(1), ..., r(C))$ does not necessarily intersect with the Pareto frontier. Figure 7.8 illustrates these situations for $C = 2$ and $C = 3$.

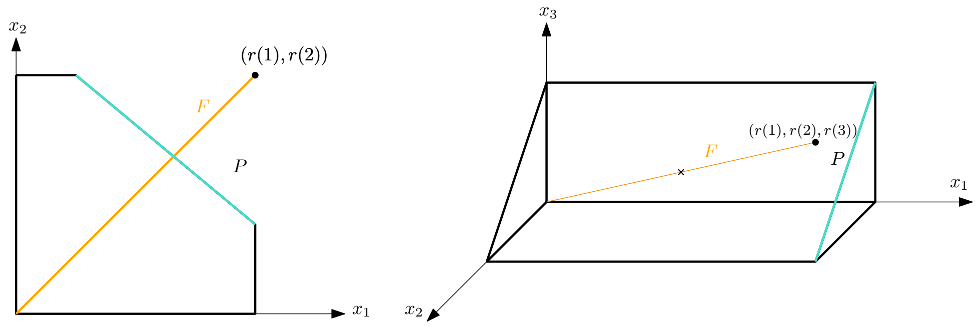


Fig. 7.8: Relation between the Pareto frontier P (light blue) and the set of opportunity fair allocations F (orange) for two and three groups. For $C = 2$ they always intersect, however it is not always the case for $C > 2$ as illustrated on the right (the cross marks the largest feasible fair matching).

7.9.7 Proof of Proposition 7.12

Proposition 7.12. For any C-colored matroid M, it holds,

$$\mathrm{PoF}(M) \leqslant \frac{1}{2} \cdot \frac{\max_{c\in[C]} \mathrm{r}(c)}{\min_{c\in[C]} \mathrm{r}(c)} + \frac{C}{4} \cdot \left(\frac{\max_{c\in[C]} \mathrm{r}(c)}{\min_{c\in[C]} \mathrm{r}(c)}\right)^2 + \frac{1}{4C} \cdot \mathbb{1}\{C \text{ odd}\}.$$

Moreover, whenever all groups have the same rank, the resulting bound is tight.

Proof. From Corollary 7.8 we know that $\mathrm{PoF}(M) = \frac{\mathrm{r}([C])}{\sum_{c\in[C]} \mathrm{r}(c)} \cdot \max_{\Lambda\subseteq[C]} \frac{\sum_{c\in\Lambda} \mathrm{r}(c)}{\mathrm{r}(\Lambda)}$. Let Λ^* be the argmax in the equation, and let us reorder the groups so $\Lambda^* = [\ell]$ for some $\ell \in [C]$. Denote $\gamma := \frac{\max_{c\in[C]} \mathrm{r}(c)}{\min_{c\in[C]} \mathrm{r}(c)}$. First, remark that the sub-addivitiy of the r function (consequence of non-negativity and submodularity) implies the following inequality,

$$\mathrm{r}([C]) - \mathrm{r}([\ell]) \leqslant \mathrm{r}([C] \setminus [\ell]) \leqslant \sum_{c=\ell+1}^{C} \mathrm{r}(c) \leqslant (C-\ell) \max\{\mathrm{r}(c), c \in [C]\}.$$

It follows that

$$\mathrm{PoF}(M) = \frac{\mathrm{r}([C])}{\mathrm{r}([\ell])} \cdot \frac{\sum_{c\in[\ell]} \mathrm{r}(c)}{\sum_{c\in[C]} \mathrm{r}(c)} \leqslant \left(1 + \frac{\mathrm{r}([C]) - \mathrm{r}([\ell])}{\mathrm{r}([\ell])}\right) \frac{\ell}{C}\gamma \leqslant (1 + (C-\ell)\gamma)\frac{\ell}{C}\gamma.$$

The right-hand side of the previous inequality is maximized (subject to $\ell \in \mathbb{N}$) at $\ell = C/2 + \mathbb{1}\{C \text{ odd}\}/2\gamma$, which leads to the stated upper bound. Concerning the general result of the tightness of the bound, consider the constructions illustrated in Figures 7.9 and 7.10 where $\lfloor \frac{C}{2} \rfloor$ groups are isolated and $\lceil \frac{C}{2} \rceil$ compete for the same resources, respectively for transversal and graphic matroids, for (left) $C = 4$ and $\mathrm{r}(c) = 3$ for all $c \in [C]$, and (right) $C = 5$ with $\mathrm{r}(c) = 5$ for all $c \in [C]$. Suppose, moreover, that all groups have rank r, for some $r \in \mathbb{N}$. An opportunity fair allocation must allocate at most $r/\lceil \frac{C}{2} \rceil$ resources to each group. It follows that

$$\mathrm{PoF} = \frac{Cr/\lceil \frac{C}{2} \rceil}{r + r\lfloor \frac{C}{2} \rfloor} = \frac{(1 + \lfloor C/2 \rfloor)\lceil C/2 \rceil}{C} = \frac{1}{2} + \frac{C}{4} + \frac{\mathbb{1}\{C \text{ odd}\}}{4C}. \qquad \square$$

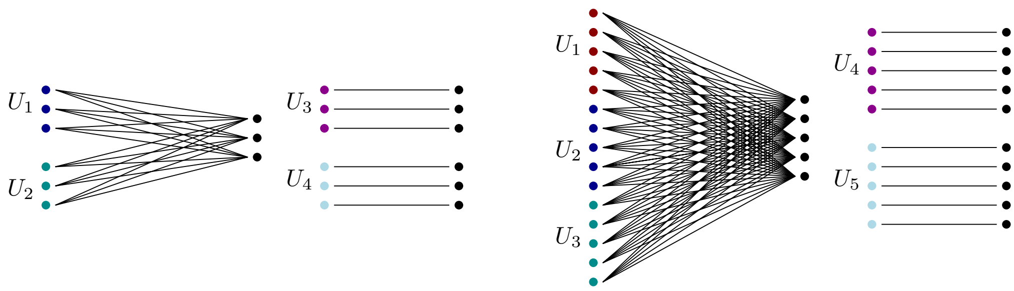


Fig. 7.9: Tight bound example for PoF as stated in Proposition 7.12 for transversal matroids with (left) four groups and (right) five groups.

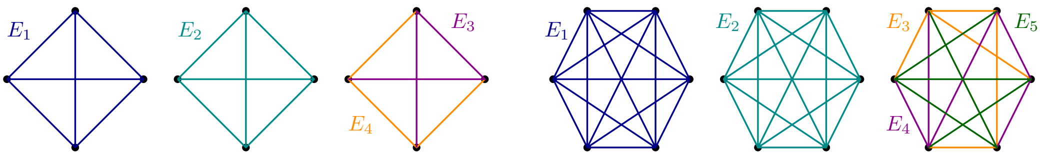


Fig. 7.10: Tight bound example for PoF as stated in Proposition 7.12 for graphic matroids with (left) four groups and (right) five groups.

7.9.8 Proof of Proposition 7.14

Proposition 7.14. Let M be a C-colored matroid. Suppose that for any c, c' in $[C]$, $\mathrm{r}(c) = \mathrm{r}(c')$. Whenever $\rho \in [1/C, 1/(C-1)]$, $\mathrm{PoF}(M) = 1$. Otherwise,

$$\mathrm{PoF}(M) \leqslant \rho \max\left(\frac{C - \lfloor C\rho \rfloor + 1}{C\rho - \lfloor C\rho \rfloor + 1}, C - \lfloor C\rho \rfloor\right) \leqslant \rho((1-\rho)C + 1).$$

In addition, the first upper bound is jointly tight in ρ and C.

Proof. Let M be a polymatroid such that all groups $c \in [C]$ have the same rank r. Suppose $\rho \in [1/C, 1/(C-1)]$ and $\mathrm{PoF}(M) > 1$. Bounding the PoF as in Theorem 7.9 plus using the fact that $\rho \leqslant 1/(C-1)$ leads to,

$$\mathrm{PoF}(M) \leqslant \rho(C-1) \leqslant 1,$$

which is a contradiction.

Suppose $\rho \in [1/(C-1), 1]$. Let $\alpha^* = \max\{\alpha \in [0,1] : \alpha(\mathrm{r}, ..., \mathrm{r}) \in M\}$, the price of fairness is written as,

$$\mathrm{PoF} = \frac{\mathrm{r}([C])}{\alpha^* \sum_{c\in[C]} \mathrm{r}} = \frac{\rho \sum_{c\in[C]} \mathrm{r}}{\alpha^* \sum_{c\in[C]} \mathrm{r}} = \frac{\rho}{\alpha^*}.$$

Therefore, the stated upper bound comes from proving that

$$\frac{1}{\alpha^*} \leqslant \max\left\{\frac{C - \lfloor C\rho \rfloor + 1}{C\rho - \lfloor C\rho \rfloor + 1}, C - \lfloor C\rho \rfloor\right\}. \tag{7.5}$$

Let $\sigma \in \Sigma([C])$ be a permutation, $c \in [C]$, and denote, for $\ell \in [C]$,

$$\alpha_\ell(\sigma) := \frac{\mathrm{r}(\sigma([\ell]))}{\sum_{t\in[\ell]} \mathrm{r}(\sigma(t))},$$

where $\mathrm{r}(\sigma([\ell])) = \mathrm{r}(\{\sigma(1), ..., \sigma(\ell)\})$ corresponds to the size of a maximum size allocation in the submatroid obtained by restricting to the groups in the first ℓ entries of σ. With this in mind, it follows,

$$\alpha^* = \min_{\sigma\in\Sigma([C])} \min_{\ell\in[C]} \alpha_\ell(\sigma).$$

Therefore, in order to prove Equation (7.5) it is enough to prove that for any permutation $\sigma \in \Sigma([C])$ and any $\ell \in [C]$, Equation (7.5) holds for $\alpha_\ell(\sigma)$. Let us prove the property for $\sigma = \mathrm{I}_C$, the identity permutation. Notice this is done without loss of generality as the same argument will work for any other permutation σ. It follows,

$$\begin{aligned}
\alpha_\ell &= \frac{\mathrm{r}([\ell]) + \sum_{t\in[\ell]} \mathrm{r}(t) - \sum_{t\in[\ell]} \mathrm{r}(t)}{\sum_{t\in[\ell]} \mathrm{r}(t)} \\
&= 1 - \frac{\sum_{t\in[\ell]} \mathrm{r}(t) - \mathrm{r}([\ell])}{\sum_{t\in[\ell]} \mathrm{r}(t)} \\
&= 1 - \frac{\sum_{t\in[L]} \mathrm{r}(t) - \sum_{t=\ell+1}^{C} \mathrm{r}(t) - \mathrm{r}([\ell])}{\sum_{t\in[\ell]} \mathrm{r}(t)}
\end{aligned}$$

$$
\begin{aligned}
&= 1 - \frac{C\mathrm{r} - \sum_{t=\ell+1}^{C} \mathrm{r}(t) - \mathrm{r}([\ell])}{\ell \mathrm{r}} \\
&= 1 - \frac{C\mathrm{r} - \mathrm{r}([C]) - \sum_{t=\ell+1}^{C} \mathrm{r}(t) - \mathrm{r}([\ell]) + \mathrm{r}([C])}{\ell \mathrm{r}} \\
&= 1 - \frac{C\mathrm{r} - \rho C\mathrm{r}}{\ell \mathrm{r}} + \frac{\sum_{t=\ell+1}^{C} \mathrm{r}(t) + \mathrm{r}([\ell]) - \mathrm{r}([C])}{\ell \mathrm{r}} \\
&= 1 - \frac{C(1-\rho)}{\ell} + \frac{\sum_{t=\ell+1}^{C} \mathrm{r}(t) + \sum_{t=\ell+1}^{C} \mathrm{r}([t-1]) - \mathrm{r}([t])}{\ell \mathrm{r}} \\
&= 1 - \frac{C(1-\rho)}{\ell} + \frac{\sum_{t=\ell+1}^{C} [\mathrm{r}(t) - \mathrm{r}([t]) + \mathrm{r}([t-1])]}{\ell \mathrm{r}}.
\end{aligned}
$$

The numerator of the third term satisfies,

$$
\sum_{t=\ell+1}^{C} [\mathrm{r}(t) - \mathrm{r}([t]) + \mathrm{r}([t-1])] \geqslant \max\{0, C(1-\rho) - (\ell-1)\}.
$$

Indeed, the term is always non-negative as the rank function is submodular and non-negative, therefore, $\mathrm{r}([t]) = \mathrm{r}([t-1] \cup \{t\}) \leqslant \mathrm{r}([t-1]) + \mathrm{r}(t)$. The second lower bound comes from,

$$
\begin{aligned}
\sum_{t=\ell+1}^{C} [\mathrm{r}(t) - \mathrm{r}([t]) + \mathrm{r}([t-1])] &= \sum_{t\in[C]} \mathrm{r}(t) - \mathrm{r}([C]) + \mathrm{r}([\ell]) - \sum_{t\in[\ell]} \mathrm{r}(t) \\
&= C\mathrm{r}(1-\rho) + \mathrm{r}([\ell]) - \ell r \\
&\geqslant C\mathrm{r}(1-\rho) - (\ell-1)\mathrm{r},
\end{aligned}
$$

where we have used that $\mathrm{r}([\ell]) \geqslant \mathrm{r}(\ell) = \mathrm{r}$. It follows,

$$
\alpha_\ell \geqslant 1 - \frac{C(1-\rho)}{\ell} + \max\left\{0, \frac{C(1-\rho) - (\ell-1)}{\ell}\right\} = \max\left\{\frac{C(\rho-1)+\ell}{\ell}, \frac{1}{\ell}\right\}.
$$

In particular, as the lower bound over α_ℓ does not depend on the chosen permutation,

$$
\alpha^* \geqslant \min_{\ell\in[C]} \max\left\{\frac{C(\rho-1)+\ell}{\ell}, \frac{1}{\ell}\right\},
$$

whose minimum is attained at $\ell^* = (1-\rho)C + 1$. Remark the second upper bound is obtained by replacing ℓ^* in the previous inequality. Regarding the first upper bound, as ℓ must be an integer, the minimum is either reached at $\lfloor \ell^* \rfloor$ or $\lceil \ell^* \rceil$. It follows,

$$
\alpha^* \geqslant \min\left\{\frac{C(\rho-1) + \lceil (1-\rho)C+1 \rceil}{\lceil (1-\rho)C+1 \rceil}, \frac{1}{\lfloor (1-\rho)C+1 \rfloor}\right\} = \min\left\{\frac{C\rho - \lfloor C\rho \rfloor + 1}{C - \lfloor C\rho \rfloor + 1}, \frac{1}{C - \lfloor C\rho \rfloor}\right\},
$$

which concludes the proof of the stated upper bound.

Regarding the tightness of the bound, we provide the example for transversal matroids. A similar construction can be done for graphic matroids by using the Hamiltonian cycle decomposition. Let $\rho \in [1/C, 1]$, $\rho \in \mathbb{Q}$, $\mathrm{r} \gg 1$, and denote $\ell^* := \lfloor C\rho \rfloor$ and $\mathrm{r}_1 = (C\rho - \ell^*)\mathrm{r}$. We take r such that $\mathrm{r}_1 \in \mathbb{N}$. Consider the following bipartite graph,

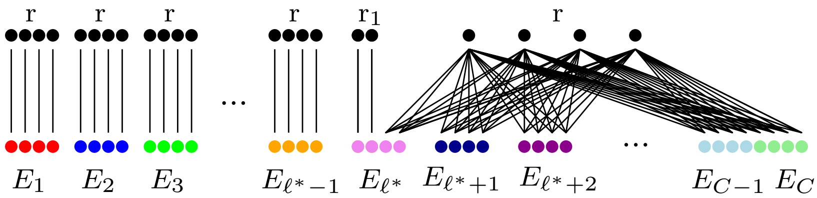


Fig. 7.11: Tight bound example for PoF as stated in Proposition 7.14 for transversal matroids.

where each group E_ℓ has r elements. All groups $\ell \in \{1, ..., \ell^* - 1\}$ are independent and have $\mathrm{r}(\ell) = \mathrm{r}$. All groups $\ell \in \{\ell^* + 1, ..., C\}$ share resources and have $\mathrm{r}(\ell) = \mathrm{r}$. Finally, group E_{ℓ^*} is a *semi-independent* group, where r_1 agents are connected to r_1 resources and $\mathrm{r} - \mathrm{r}_1$ agents belong to the clique. We obtain $\mathrm{r}(\ell^*) = \mathrm{r}$ as well. Finally, remark $\mathrm{r}([C]) = (\ell^* - 1)\mathrm{r} + \mathrm{r}_1 + \mathrm{r} = C\rho\mathrm{r}$.

We focus next on the maximum size opportunity fair allocation. Notice that, as all groups have the same rank, an allocation x is opportunity fair if $x_\ell = x_k$ for all $\ell, k \in [C]$. Since all groups $\ell \in \{\ell^* + 1, ..., C\}$ share all their resources, the highest share than can be fairly allocated to them is,

$$x_\ell = \frac{\mathrm{r}}{C - \ell^*}.$$

This allocation is feasible if and only if the remaining available resources to be allocated to ℓ^* are enough to fulfill its demand, i.e, if and only if,

$$\mathrm{r}_1 \geqslant \frac{\mathrm{r}}{C - \ell^*} \tag{7.6}$$

Moreover, remark that depending on whether Equation (7.6) holds or not, the maximum on the stated upper bound gets a different value,

$$\begin{aligned}
&\frac{\mathrm{r}}{C - \lfloor C\rho \rfloor} \leqslant (L\rho - \lfloor C\rho \rfloor)\mathrm{r} \\
\Longleftrightarrow &\frac{1}{C - \lfloor C\rho \rfloor} + 1 \leqslant C\rho - \lfloor C\rho \rfloor + 1 \\
\Longleftrightarrow &\frac{C - \lfloor C\rho \rfloor + 1}{C - \lfloor C\rho \rfloor} \leqslant C\rho - \lfloor C\rho \rfloor + 1
\end{aligned}$$

$$\iff \frac{C - \lfloor C\rho \rfloor + 1}{C\rho - \lfloor C\rho \rfloor + 1} \leqslant C - \lfloor C\rho \rfloor.$$

Suppose Equation (7.6) holds. It follows the allocation is feasible and,

$$\|x\|_1 = \frac{Cr}{C - \ell^*} = \frac{Cr}{C - \lfloor C\rho \rfloor},$$

and the Price of Fairness is equal to,

$$\mathrm{PoF} = \frac{C\rho \mathrm{r}}{\frac{C\mathrm{r}}{C - \lfloor C\rho \rfloor}} = \rho(C - \lfloor C\rho \rfloor),$$

which is indeed equal to the upper bound. Suppose Equation (7.6) does not hold. In particular, the opportunity fair allocation must allocate some share of the r resources on the clique to E_{ℓ^*}. Let $s \in (0,1)$ denote the share. We obtain the following system,

$$s\mathrm{r} + \mathrm{r}_1 = \frac{(1-s)\mathrm{r}}{C - \ell^*},$$

whose solution is given by

$$s^* = \frac{1 - (C\rho - \ell^*)(C - \ell^*)}{C - \ell^* + 1}.$$

It follows the opportunity fair allocation x has size,

$$\|x\|_1 = C \cdot \frac{(1-s^*)\mathrm{r}}{C - \ell^*} = \frac{C(C\rho - \ell^* + 1)\mathrm{r}}{C - \ell^* + 1},$$

which yields,

$$\mathrm{PoF} = \frac{C\rho \mathrm{r}}{\frac{C(C\rho - \ell^* + 1)\mathrm{r}}{C - \ell^* + 1}} = \rho \cdot \frac{C - \ell^* + 1}{C\rho - \ell^* + 1} = \rho \cdot \frac{C - \lfloor C\rho \rfloor + 1}{C\rho - \lfloor C\rho \rfloor + 1},$$

which corresponds to the stated upper bound when Equation (7.6) does not hold. □

7.9.9 Proof of Proposition 7.15

We recall that $(M_n(p))_{n\in\mathbb{N}}$ is a sequence of C-colored matroids over sets $(E^{(n)})$ such that $|E^{(n)}| = n$, randomly colored according to $p = (p_1, p_2, ..., p_C)$, and $(\mathrm{r}_n)_{n\in\mathbb{N}}$ the associated sequence of rank functions. For each $c \in [C]$, suppose the following limit exists,

$$R(p_c) := \lim_{n\to\infty} \frac{\mathbb{E}_{p_c}[\mathrm{r}_n(c)]}{n}, \tag{7.7}$$

and recall its natural extension to any subset $\Lambda \subseteq [C]$,

$$R\Big(\sum\nolimits_{c\in\Lambda} p_c\Big) := \lim_{n\to\infty} \frac{1}{n} \cdot \mathbb{E}_{(p_c)_{c\in\Lambda}}\big[\mathrm{r}_n\big(\bigcup\nolimits_{c\in\Lambda} E_c\big)\big].$$

Proposition 7.15. If $R(1) > 0$, it follows

$$\mathrm{PoF}(M_n(p)) \xrightarrow[n\to\infty]{\mathrm{P}} \max_{\Lambda\subseteq[C]} \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \cdot \frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)}, \tag{7.8}$$

where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability. Moreover, the function R is such that $R(0) = 0$, R is concave, non-decreasing, and 1-Lipschitz. Finally, for any such function R, there exists a double sequence of C-colored matroids $(M'_{n,m})$ such that, for R'_m defined similarly to R for the sequence $(M'_{n,m})_{n\in\mathbb{N}}$,

$$\|R - R'_m\|_\infty \xrightarrow[m\to\infty]{} 0.$$

Proof. Let $\mathcal{R} := \{f : [0,1] \to \mathbb{R} \mid f$ is a concave, non-decreasing, 1-Lipschitz function, and $f(0) = 0\}$. Remark that $\mathcal{R}$ is closed and convex. We divide the proof in several steps. First, we prove the function R defined in Equation (7.7) belongs to $\mathcal{R}$. Second, we prove the PoF converges in probability to the stated limit. Third, we construct a double sequence of colored matroids $(M'_{n,m})_{n,m\in\mathbb{N}}$ such that R is well approximated by R_m, where each R_m is defined as in Equation (7.7) for $(M'_{n,m})_{n\in\mathbb{N}}$.

1. Function R belongs to $\mathcal{R}$. Since $\mathcal{R}$ is closed, it is enough to prove that for each $n \in \mathbb{N}$, the mapping

$$\begin{aligned} Q_n : [0,1] &\to \mathbb{R}_+ \\ p_c &\mapsto \frac{\mathbb{E}_{p_c}[\mathrm{r}_n(c)]}{n} \end{aligned}$$

belongs to $\mathcal{R}$. Clearly, $Q_n(0) = 0$. Regarding concavity and monotonicity, remark

$$\mathbb{E}_{p_c}[\mathrm{r}_n(c)] = \sum_{S\subseteq E} \mathrm{r}_n(S)\mathbb{P}(E_c = S) = \sum_{S\subseteq E} \mathrm{r}_n(S) p_c^{|S|}(1-p_c)^{|E|-|S|},$$

which can be seen as the multi-linear extension of the rank function r_n of M_n evaluated at $(p_c, \ldots, p_c)$. It follows that $\mathbb{E}_{p_c}[\mathrm{r}_n(c)]$ is a concave and non-decreasing function as r_n is submodular [Căl+11]. Moreover, Q_n is also concave and non-decreasing. Finally, remark,

$$\mathbb{E}_{p_c}[r_n(c)] \leqslant \mathbb{E}_{p_c}[|E_c|] = np_c,$$

where n is the total number of agents in E. In particular, Q_n is 1-Lipschitz[4].

2. Convergence of PoF. To prove the convergence of the PoF, remark first that $\mathrm{r}_n(c)$ concentrates around its mean $\mathbb{E}_{p_c}[\mathrm{r}_n(c)]$. Indeed, $\mathrm{r}_n(c)$ is a function on the indicator variables $\mathbb{1}[e \in E_c]$ for $e \in E$, which are i.i.d. according to Ber(p_c). In particular, $\mathrm{r}_n(c)$ has a bounded difference of 1, as for any of the indicator variables that changes of value, the rank modifies at most in 1. The McDiarmid concentration inequality implies that,

$$\mathbb{P}\left(|\mathrm{r}_n(c) - \mathbb{E}_{p_c}[\mathrm{r}_n(c)]| \geqslant \sqrt{n\log(n)}\right) \leqslant \exp\left(\frac{-2n\log(n)}{n}\right) = \frac{1}{n^2}.$$

Added to the union bound, we obtain that,

$$\left|\sum\nolimits_{c\in[C]} \frac{\mathrm{r}_n(c)}{n} - \sum\nolimits_{c\in[C]} \frac{\mathbb{E}_{p_c}[\mathrm{r}_n(c)]}{n}\right| \xrightarrow[n\to\infty]{P} 0,$$

in other words,

$$\lim_{n\to\infty} \sum\nolimits_{c\in[C]} \frac{\mathrm{r}_n(c)}{n} = \sum\nolimits_{c\in[C]} R(p_c).$$

For any $c \in [C]$, notice that,

$$R(1) = \lim_{n\to\infty} \frac{\mathbb{E}_{p_c=1}[\mathrm{r}_n(c)]}{n} = \lim_{n\to\infty} \frac{\mathbb{E}[\mathrm{r}_n(E)]}{n} = \lim_{n\to\infty} \frac{\mathrm{r}_n(E)}{n}.$$

Next, since R is concave,

$$\lim_{n\to\infty} \frac{\mathbb{E}_{p_c}[\mathrm{r}_n(c)]}{n} = R(p_c) = R(p_c \cdot 1 + (1-p_c)\cdot 0) \geqslant p_c R(1) + (1-p_c)R(0) = p_c \lim_{n\to\infty} \frac{\mathrm{r}_n(E)}{n}.$$

Since $R(1) > 0$, we obtain that both $\mathrm{r}_n(E) = \Omega(n)$ and $\mathbb{E}_{p_c}[\mathrm{r}_n(c)] = \Omega(n)$. Putting all together, we conclude the following,

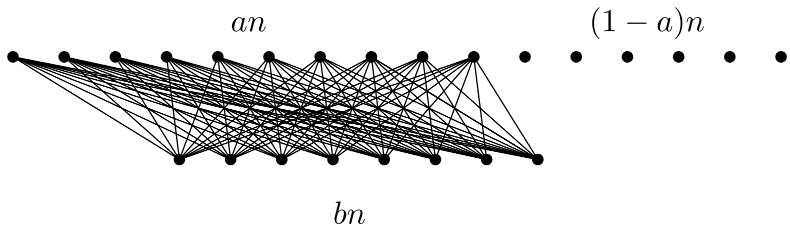
$$\mathrm{PoF}(M_n) = \max_{\Lambda\subseteq[C]} \frac{\mathrm{r}_n([C])}{\sum_{c\in[C]} \mathrm{r}_n(c)} \cdot \frac{\sum_{c\in\Lambda} \mathrm{r}_n(c)}{\mathrm{r}_n(\Lambda)} \xrightarrow[n\to\infty]{} \max_{\Lambda\subseteq[C]} \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \cdot \frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)},$$

where we used the assumption that R exists and that $R(p_c) > 0$ for all c.

3. Approximation result. To approximate the functions in $\mathcal{R}$, we will construct a family of matroids able to produce a family of piece-wise functions $f \in \mathcal{R}$ whose convex hull i dense on the set $\mathcal{R}$. Let $0 \leqslant b \leqslant a \leqslant 1$ be two real values and $n \in \mathbb{N}$, such that an, bn, and $(1-a)n$ are integer values. Consider the following graph

[4]Remark the function is defined over the interval $[0,1]$. Therefore, Q_n is concave, increasing, $Q(0) = 0$, and $Q(x) \leqslant x$ for any $x \in [0,1]$, if and only if the function is 1-Lipschitz.

containing a complete bipartite graph with sides of sizes an and bn, respectively, and $(1-a)n$ isolated vertices, as in the figure below,



Let M_n be the associated transversal matroid. Given a random coloring according to a vector $p = (p_1, ..., p_C)$, notice that, for n large enough,

$$\mathbb{E}_{p_c}[\mathrm{r}_n(c)] = \min\{ap_c n, bn\}.$$

For this sequence of matroids, it follows that,

$$R(p_c) = \lim_{n\to\infty} \frac{\mathbb{E}_{p_c}[\mathrm{r}_n(c)]}{n} = \min\{ap_c, b\}.$$

We denote

$$\mathcal{T} := \{f : [0,1] \to \mathbb{R} \mid \exists a, b \in \mathbb{R}_+, f(t) = \min\{at, b\}, \forall t \in [0,1]\}.$$

In particular, all functions in $\mathcal{T}$ can be obtained by the previous construction. Consider next the set,

$$\mathcal{H} := \{f : [0,1] \to \mathbb{R} \mid f \text{ is piece-wise linear, concave, non-decreasing, 1-Lipschitz, and } f(0) = 0\}.$$

We claim that any function in $\mathcal{H}$ can be obtained as convex combinations of functions within $\mathcal{T}$. Indeed, for $f \in \mathcal{H}$ consisting in two pieces of value a and then $b \leqslant a$, i.e, such that there exists $t^* \in [0,1]$,

$$f(t) = \begin{cases} at & t \leqslant t^*, \\ b(t-t^*) + at^* & t \geqslant t^*, \end{cases}$$

it is enough to take

$$f_1 : \begin{array}{l}[0,1] \to \mathbb{R}, \\ t \mapsto at\end{array} \quad f_2 : \begin{array}{l}[0,1] \to \mathbb{R} \\ t \mapsto \max(at, at^*)\end{array}$$

as $f \equiv \frac{b}{a} f_1 + (1 - \frac{b}{a}) f_2$. For the rest of functions within $\mathcal{H}$, the construction is done inductively. Consider $f \in \mathcal{H}$ to be $(m+1)$-piece-wise linear, for $m \geqslant 2$. Let

$0 \leqslant c \leqslant b \leqslant a \leqslant 1$ be the last three linear slopes of f with respective changes at $t_1 \leqslant t_2$, as illustrated in Figure 7.12.

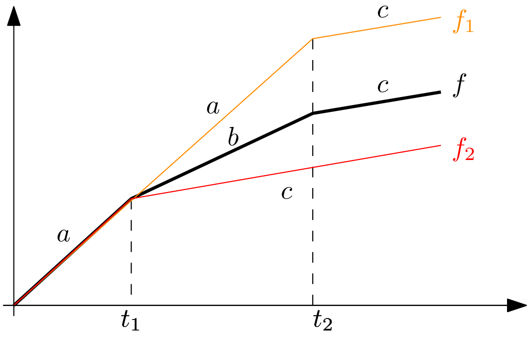


Fig. 7.12: Example piece-wise function

Consider next,

$$f_1(t) := \begin{cases} f(t) & t \leqslant t_1 \\ f(t_1) + at & t \in [t_1, t_2] \\ f(t_1) + a(t_2 - t_1) + ct & t \geqslant t_2 \end{cases} \quad \text{and} \quad f_2(t) := \begin{cases} f(t) & t \leqslant t_1 \\ f(t_1) + ct & t \geqslant t_1 \end{cases}$$

Remark both f_1 and f_2 are m-piece-wise linear. It is not hard to check that

$$f \equiv \left(\frac{b-c}{a-c}\right) f_1 + \left(1 - \left(\frac{b-c}{a-c}\right)\right) f_2.$$

We conclude the proof by showing that $\mathcal{H}$ is dense in $\mathcal{R}$. Let $R \in \mathcal{R}$ be fixed. For $m \in \mathbb{N}$, divide the interval $[0, 1]$ in m pieces $\{0, \frac{1}{m}, \frac{2}{m}, \dots, \frac{m-1}{m}, 1\}$, and define the m-piece-wise linear function that interpolates R, as it follows,

$$f(t) = R(i) + t\left(R\left(\frac{i+1}{m}\right) - R\left(\frac{i}{m}\right)\right), \text{ for } t \in \left[\frac{i}{m}, \frac{i+1}{m}\right], \text{ for } i \in \{0, 1, \dots, m\}.$$

For $t \in [\frac{i}{m}, \frac{i+1}{m}]$, by monotonicity of R and f, it follows,

$$\begin{aligned} |R(t) - f(t)| &\leqslant \max\left\{R\left(\frac{i+1}{m}\right) - f\left(\frac{i}{m}\right); f\left(\frac{i+1}{m}\right) - R\left(\frac{i}{m}\right)\right\} \\ &= R\left(\frac{i+1}{m}\right) - R\left(\frac{i}{m}\right) \\ &\leqslant \frac{1}{m} \xrightarrow[m\to\infty]{} 0, \end{aligned}$$

where the last inequality comes from the fact that R is 1-Lipschitz. In particular, $\|R - f\|_\infty \to 0$. $\square$

7.9.10 Proof of Theorem 7.16

Theorem 7.16. Let $\pi \in [0,1]$ be fixed. Consider the sets,

$$\mathcal{R} := \{f : [0,1] \to \mathbb{R} \mid f \text{ is a concave, non-decreasing, 1-Lipschitz function, and } f(0) = 0\},$$
$$\Delta_\pi^C := \{p \in \Delta^C \mid \max_{c\in[C]} p_c = \pi\}.$$

It follows,

$$\max_{p\in\Delta_\pi^C} \max_{R\in\mathcal{R}} \max_{\Lambda\subseteq[C]} \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \cdot \frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)} = \max_{\lambda\in[C]} \psi_\lambda\left(\frac{1-(C-\lambda)\pi}{C}\right) \leqslant C - \frac{1}{\pi}, \tag{7.9}$$

where $\psi_\lambda : [-\lambda, \frac{1}{C}] \to \mathbb{R}$, for each $\lambda \in [C]$, is given by,

$$\psi_\lambda(q) = \begin{cases} \lambda & q \in [-\lambda, 0], \\ \frac{\lambda}{(\lambda C q-1)^2} \cdot \left(1+q(1-2\lambda)+C(\lambda-2+\lambda q)q-2\sqrt{(\lambda-1)(C-1)(1-Cq)(1-\lambda q)q}\right) & q \in \left(0, \frac{(\lambda-1)}{\lambda(C-1)}\right], \\ 1 & q \in \left(\frac{(\lambda-1)}{\lambda(C-1)}, \frac{1}{C}\right]. \end{cases} \tag{7.10}$$

The proof of Theorem 7.16 consists on constructing an optimal solution of the triple optimization problem in Equation (7.9) by starting from an instance (p_0, Λ_0, R_0) and iteratively modifying p, Λ, and R. Before giving the formal proof, we show some useful technical lemmas. We define

$$\begin{aligned} F : \Delta^C \times \mathcal{R} \times 2^{[C]} &\longrightarrow [1,\infty) \\ (p, R, \Lambda) &\longrightarrow F(p,R,\Lambda) := \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \cdot \frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)}. \end{aligned}$$

Remark that whenever $|\Lambda| \in \{1, C\}$, $F(p,R,\Lambda) = 1$, for any $p, R \in \Delta^C \times \mathcal{R}$. Indeed,

$$F(p,R,\{\bar{c}\}) = \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \leqslant 1,$$

where the inequality comes from the concavity of R and the fact that $\sum_{c\in[C]} p_c = 1$, so $\sum_{c\in[C]} R(p_c) \leqslant R(\sum_{c\in[C]} p_c)$. Similarly,

$$F(p,R,[C]) = \frac{R(1)}{R(\sum_{c\in[C]} p_c)} = 1.$$

Therefore, from now on, **we suppose $\mathbf{1 < |\Lambda| < C}$**. The function F is invariant to scaling R by non-null constants, i.e., $F(p,R,\Lambda) = F(p,\alpha R,\Lambda)$ for any $\alpha \neq 0$. In addition, F evaluates R at $C+2$ points: $(p_c)_{c\in[C]}$, $\sum_{c\in\Lambda} p_c$, and 1. Since,

$$F(p,R,\Lambda) = \frac{R(1)}{R(\sum_{c\in\Lambda} p_c)}\left(1 - \frac{\sum_{c\in[C]\setminus\Lambda} R(p_c)}{\sum_{c\in[C]} R(p_c)}\right),$$

F is decreasing on $R(\sum_{c\in\Lambda} p_c)$ and $(R(p_c))_{c\in[C]\setminus\Lambda}$ and increasing on $R(1)$ and $(R(p_c))_{c\in\Lambda}$. Figure 7.13 illustrates a function $R \in \mathcal{R}$ for $C = 5$ and $\Lambda = \{1,2,4\}$, with the red dots indicating the values where F is decreasing, and the blue dots those where F is increasing.

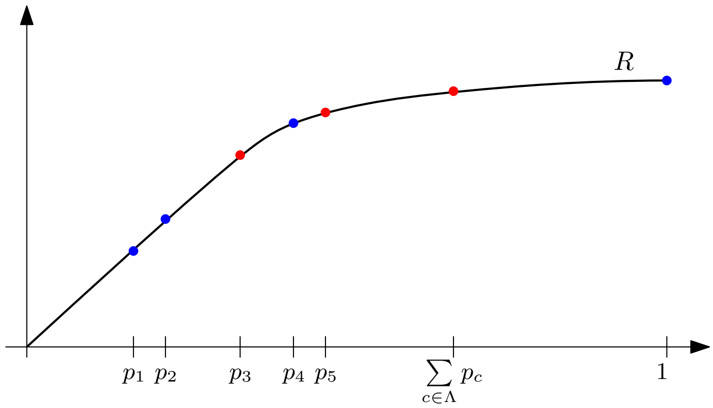


Fig. 7.13: Increasing (blue) and decreasing (red) points for function F

The construction in the proof of Theorem 7.16 will be done by playing with both: the position (over the horizontal axis) of the red and blue dots and their values.

Lemma 7.20. *Given $(p,R,\Lambda) \in \Delta^C \times \mathcal{R} \times 2^{[C]}$, we can always construct $R' \in \mathcal{R}$ such that either $F(p,R',\Lambda) > F(p,R,\Lambda)$ or $R' = R$.*

Proof. Given $(p,R,\Lambda) \in \Delta^C \times \mathcal{R} \times 2^{[C]}$, it is enough with picking $R' \in \mathcal{R}$ satisfying,

$$\begin{aligned}
&R'\left(\sum\nolimits_{c\in\Lambda} p_c\right) \leqslant R\left(\sum\nolimits_{c\in\Lambda} p_c\right)\\
&R'(p_c) \leqslant R(p_c), \forall c \in [C]\setminus\Lambda\\
&R(1) \leqslant R'(1)\\
&R(p_c) \leqslant R'(p_c), \forall c \in \Lambda.
\end{aligned}$$

For example, suppose $C = 3$, $0 < p_1 < p_2 < p_3 < p_2 + p_3 < 1$, and $\Lambda = \{2, 3\}$. Starting from R we can take $R' \in \mathcal{R}$ such that

$$R'(x) = \begin{cases} x \cdot \frac{R(p_2)}{p_2} & x \in [0, p_2] \\ R(x) & x \in [p_2, p_3] \\ R(p_3) + (x - p_3) \cdot \frac{R(1) - R(p_3)}{1 - p_3} & x \in [p_3, 1] \end{cases}$$

as illustrated in Figure 7.14,

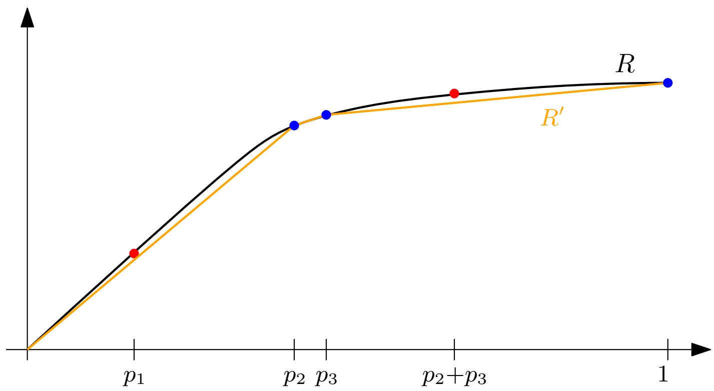


Fig. 7.14: Function R'

Remark that

$$\begin{aligned} R'(p_2 + p_3) &= R(p_3) + p_2 \cdot \frac{R(1) - R(p_3)}{1 - p_3} \\ &= \frac{p_2}{1 - p_3} \cdot R(1) + \left(1 - \frac{p_2}{1 - p_3}\right) \cdot R(p_3) \\ &\leqslant R\left(\frac{p_2}{1 - p_3} + \left(1 - \frac{p_2}{1 - p_3}\right) \cdot p_3\right) \\ &= R\left(\frac{1}{1 - p_3} \cdot p_2(1 - p_3) + p_3\right) = R(p_2 + p_3), \end{aligned}$$

where the inequality comes from R's concavity. Similarly,

$$\begin{aligned} R'(p_1) &= \frac{p_1}{p_2} \cdot R(p_2) \\ &= \frac{p_1}{p_2} \cdot R(p_2) + \left(1 - \frac{p_1}{p_2}\right) R(0) \\ &\leqslant R\left(\frac{p_1}{p_2} \cdot p_2 + \left(1 - \frac{p_1}{p_2}\right) \cdot 0\right) = R(p_1), \end{aligned}$$

where we have used R's concavity and that $R(0) = 0$. Finally, remark $R(1) = R'(1)$ and $R'(p_c) = R(p_c)$ for $c \in \Lambda$. $\square$

Lemma 7.21. *Let $\pi \in [0, 1]$ be fixed, $(p, R, \Lambda) \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$, and $c^* = \arg\max_{c \in \Lambda} p_c$ (if several c^* exists, pick one at random). Consider $p', p'' \in \Delta_\pi^C$ given by,*

$$\forall c \in [C], p'_c = \begin{cases} p_c & c \in [C] \setminus \Lambda \text{ or } c = c^* \\ \frac{1}{|\Lambda|-1} \sum_{c \in \Lambda \setminus \{c^*\}} p_c & c \in \Lambda \setminus \{c^*\}, \end{cases}$$

$$\forall c \in [C], p''_c = \begin{cases} p_c & c \in [C] \setminus \Lambda \\ \frac{1}{|\Lambda|} \sum_{c \in \Lambda} p_c & c \in \Lambda. \end{cases}$$

It follows $F(p, R, \Lambda) \leqslant F(p', R, \Lambda)$ and $F(p, R, \Lambda) \leqslant F(p'', R, \Lambda)$.

To prove Lemma 7.21 we introduce the following definition.

Definition 7.22. For $x \in \mathbb{R}_+^C$ a vector, we denote $x_{(c)}$ to its c-th highest entry. Given $x, y \in \mathbb{R}_+^C$, we say that x **majorizes** y if

$$\sum_{c=1}^{\lambda} x_{(c)} \geqslant \sum_{c=1}^{\lambda} y_{(c)}, \text{ for all } \lambda \in [C], \text{ and } \sum_{c=1}^{C} x_c = \sum_{c=1}^{C} y_c.$$

In addition, we state **Kamarata's inequality**: Let $x, y \in \mathbb{R}^C$ be two vectors such that x majorizes y. For any concave function f, it follows,

$$\sum_{c \in [C]} f(x_c) \leqslant \sum_{c \in [C]} f(y_c).$$

Proof of Lemma 7.21. We prove the stated result for p'. For p'' the argument is analogous. Recall that F is increasing on $\sum_{c \in \Lambda} R(p_c)$. We prove that $(p_c)_{c \in \Lambda}$ majorizes $(p'_c)_{c \in \Lambda}$ and conclude by using Karamata's inequality over R. First, remark

$$\sum_{c \in \Lambda} p'_c = p_{c^*} + \sum_{c \in \Lambda \setminus \{c^*\}} \frac{1}{|\Lambda| - 1} \sum_{c \in \Lambda \setminus \{c^*\}} p_c = p_c^* + \sum_{c \in \Lambda \setminus \{c^*\}} p_c = \sum_{c \in \Lambda} p_c.$$

Regarding the inequality, assume without loss of generality that $\Lambda = \{1, ..., m\}$ and $p_1 \geqslant p_2 \geqslant ... \geqslant p_m$. In particular, notice $p'_1 \geqslant p_1$ (as they are equal). For any $\lambda \in \{2, ..., m-1\}$, it follows,

$$\sum_{c=1}^{\lambda} p'_c = p_1 + \sum_{c=2}^{\lambda} \frac{1}{m-1} \sum_{c=2}^{m} p_c$$

$$
\begin{aligned}
&= p_1 + \frac{\lambda-1}{m-1}\sum_{c=2}^{m} p_c \\
&= p_1 + \frac{\lambda-1}{m-1}\sum_{c=2}^{\lambda} p_c + \frac{\lambda-1}{m-1}\sum_{c=\lambda+1}^{m} p_c \\
&= p_1 + \sum_{c=2}^{\lambda} p_c - \frac{m-\lambda}{m-1}\sum_{c=2}^{\lambda} p_c + \frac{\lambda-1}{m-1}\sum_{c=\lambda+1}^{m} p_c \\
&\leqslant p_1 + \sum_{c=2}^{\lambda} p_c = \sum_{c=1}^{\lambda} p_c,
\end{aligned}
$$

where the last inequality comes from the fact that,

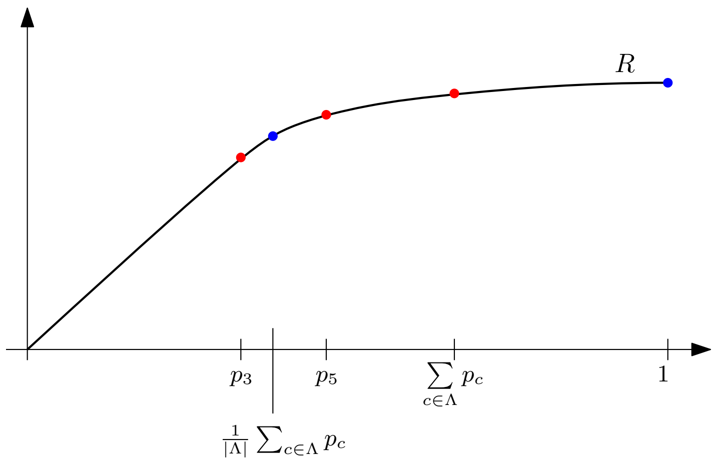
$$(m-\lambda)(\lambda-1)p_\lambda \leqslant (m-\lambda)\cdot\sum_{c=2}^{\lambda} p_c \text{ and } (\lambda-1)\cdot\sum_{c=\lambda+1}^{m} p_c \leqslant (\lambda-1)(m-\lambda)p_{\lambda+1},$$

and therefore,

$$\frac{\lambda-1}{m-1}\sum_{c=\lambda+1}^{m} p_c - \frac{m-\lambda}{m-1}\sum_{c=2}^{\lambda} p_c \leqslant \frac{(m-\lambda)(\lambda-1)}{m-1}\cdot(p_{\lambda+1}-p_\lambda) \leqslant 0,$$

for any $\lambda \in \{2, ..., m-1\}$. □

Lemma 7.21 allows to replace all elements $p_c \in \Lambda$ by one single value equal to their mean. In particular, Figure 7.13 becomes



The main issue with the uniformization of the probabilities within Λ is the fact that any transformation done to the vector $p \in \Delta_\pi^C$ must produce a vector within Δ_π^C, i.e., the maximum-value must remain unchanged (although it could eventually change of index). The following Lemma shows that for any instance $(p, R, \Lambda) \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$, we can always modify R and Λ such that the maximum-value entry of p stays outside of Λ, with a transformation that does not decrease the value of $F(p, R, \Lambda)$.

Lemma 7.23. *Let $\pi \in [0,1]$ be fixed, $(p, R, \Lambda) \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$, and $c^* = \arg\max_{c\in[C]} p_c$ (if several c^* exists, pick one at random), i.e., $p_{c^*} = \pi$. Then, we can always construct $(R', \Lambda') \in \mathcal{R} \times 2^{[C]}$ such that $c^* \notin \Lambda'$ and $F(p, R, \Lambda) \leqslant F(p, R', \Lambda')$.*

Proof. Suppose that $c^* \in \Lambda$. Apply the partial uniformization technique to p of Lemma 7.21 leaving p_{c^*} unchanged. Denote

$$q := \frac{1}{|\Lambda| - 1} \sum_{c \in \Lambda \setminus \{c^*\}} p_c.$$

Since $p_{c^*} \in \Lambda$, remark F is increasing at $R(1), R(q)$, and $R(\pi)$. Notice that $0 < q < \pi < \sum_{c\in\Lambda} p_c < 1$. Apply Lemma 7.20 and replace R by

$$R'(x) = \begin{cases} R(x) & x \in [0, q] \\ R(q) + (x - q)\frac{R(\pi) - R(q)}{\pi - q} & x \in [q, \pi] \\ R(\pi) + (x - \pi)\frac{R(1) - R(\pi)}{1 - \pi} & x \in [\pi, 1], \end{cases}$$

as illustrated in Figure 7.15, for $C = 5$, $\Lambda = \{1, 2, 4\}$, and $p_5 = \pi$. Remark that for $x \in [q, \pi]$ no value $R(x)$ is considered on F, which in particular allows to replace R by the linear segment between the points $(q, R(q))$ and $(\pi, R(\pi))$.

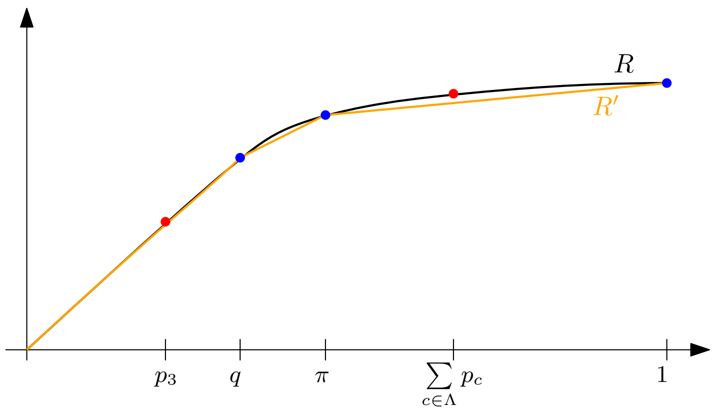


Fig. 7.15: Function R' Lemma 7.23

Next, we show we can find $\Lambda' \subseteq [C] \setminus \{c^*\}$ and $R'' \in \mathcal{R}$ starting from R' such that $F(p, R', \Lambda) \leqslant F(p, R'', \Lambda')$. Given $\varepsilon > 0$, consider

$$R'_\varepsilon(x) := \begin{cases} R'(x) & x \in [0, \pi] \\ R'(\pi) + (x - \pi)\left(\frac{R(1)-R(\pi)}{1-\pi} + \varepsilon\right) & x \in [\pi, 1]. \end{cases}$$

Let

$$\varepsilon^* = \arg\max\{\varepsilon : R'_\varepsilon \in \mathcal{R} \text{ and } F(p, R'_\varepsilon, \Lambda) \geqslant F(p, R', \Lambda)\},$$

and set $R'' = R'_{\varepsilon^*}$. We claim that

$$\frac{R(1) - R(\pi)}{1 - \pi} + \varepsilon^* = \frac{R(\pi) - R(q)}{\pi - q},$$

i.e., the segment between $(q, R''(q))$ and $(\pi, R''(\pi))$ has the same slope as the one between $(\pi, R''(\pi))$ and $(1, R''(1))$, as illustrated in Figure 7.16, for $C = 5$, $\Lambda = \{1, 2, 4\}$, and $p_5 = \pi$. Clearly, R'_{ε^*} belongs to $\mathcal{R}$. Regarding the increase on the

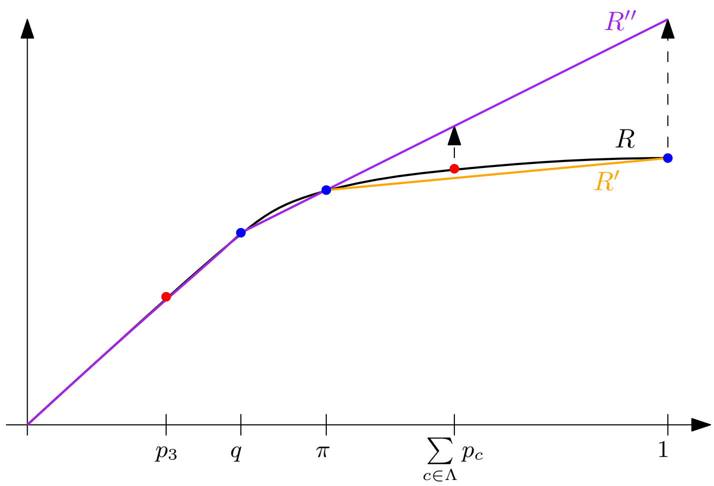


Fig. 7.16: Function R''

value of the function F, we show that the for any $\varepsilon > 0$,

$$R'_\varepsilon\left(\sum_{c\in\Lambda} p_c\right) - R'\left(\sum_{c\in\Lambda} p_c\right) \leqslant R'_\varepsilon(1) - R'(1),$$

i.e, that the increase of the blue dot in Figure 7.16 is larger than the increase of the red dot. It follows,

$$\frac{d}{d\varepsilon}\left[\frac{R'_\varepsilon(1)}{R'_\varepsilon(\sum_{c\in\Lambda} p_c)}\right] = \frac{d}{d\varepsilon}\left[\frac{R'(\pi) + (1-\pi)\left(\frac{R(1)-R(\pi)}{1-\pi} + \varepsilon\right)}{R'(\pi) + (\sum_{c\in\Lambda} p_c - \pi)\left(\frac{R(1)-R(\pi)}{1-\pi} + \varepsilon\right)}\right]$$

$$= \frac{R'(\pi)(1-\sum_{c\in\Lambda} p_c)}{\left(R'(\pi)+(\sum_{c\in\Lambda} p_c-\pi)\left(\frac{R(1)-R(\pi)}{1-\pi}+\varepsilon\right)\right)^2},$$

which is always non-negative. In particular, R'' is rewritten as

$$R''(x) := \begin{cases} R(x) & x\in[0,q] \\ R(q)+(x-q)\left(\frac{R(\pi)-R(q)}{\pi-q}\right) & x\in[q,1]. \end{cases} \tag{7.11}$$

To ease the notation, we drop the $''$ from R'' and denote $\alpha = {}^{(R(\pi)-R(q))}/_{(\pi-q)}$. Finally, we construct $\Lambda' \subseteq [C]\setminus\{c^*\}$ such that $F(p,R,\Lambda) \leqslant F(p,R,\Lambda')$. The analysis is split depending on whether a value $p_{\bar{c}}$ with $\bar{c}\in[C]\setminus\Lambda$ (a red dot) lies between q and π or not. Suppose it does. We claim that considering $\Lambda' := \Lambda\setminus\{c^*\}\cup\{\bar{c}\}$ we obtain the stated result. Indeed, although swapping the elements should decrease the value of the function F (as we obtain a higher-value red dot and a lower-value blue dot), the effect is compensated by the fact that $\sum_{c\in\Lambda'} p_c < \sum_{c\in\Lambda} p_c$. Figure 7.17 illustrates the swapping.

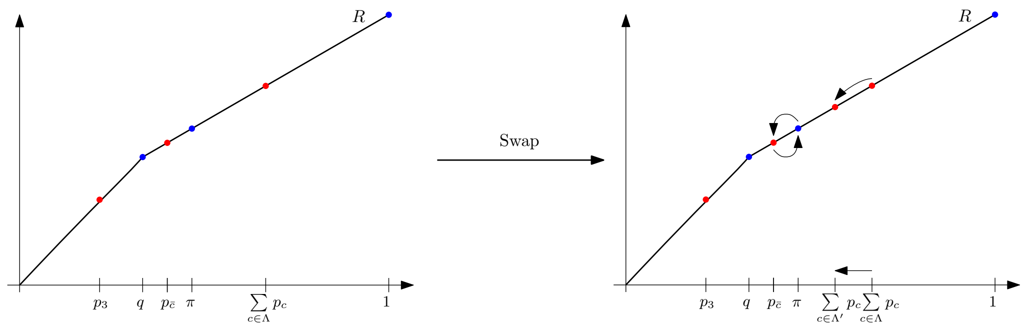


Fig. 7.17: Swapping $p_{\bar{c}}$ and π within Λ

To prove that $F(p,R,\Lambda) \leqslant F(p,R,\Lambda')$, we check that

$$\frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)} \leqslant \frac{\sum_{c\in\Lambda'} R(p_c)}{R(\sum_{c\in\Lambda'} p_c)}. \tag{7.12}$$

For $z\in[p_{\bar{c}},\pi]$, consider,

$$Q(z) := \frac{(|\Lambda|-1)R(q)+R(\pi-p_{\bar{c}}+z)}{R((|\Lambda|-1)q+\pi-p_{\bar{c}}+z)} = \frac{|\Lambda|R(q)+\alpha(\pi-p_{\bar{c}}+z-q)}{R(q)+\alpha((|\Lambda|-2)q+\pi-p_{\bar{c}}+z)}$$

where the last equality comes from using R's definition (7.11). In particular Equation (7.12) holds if and only if $Q(\pi) \leqslant Q(p_{\bar{c}})$. Notice,

$$\begin{aligned}\frac{d}{dz}Q(z) &= \frac{(R(q)+\alpha((|\Lambda|-2)q+\pi-p_{\bar{c}}+z))\alpha-(|\Lambda|R(q)+\alpha(\pi-p_{\bar{c}}+z-q))\alpha}{\left[R(q)+\alpha((|\Lambda|-2)q+\pi-p_{\bar{c}}+z)\right]^2}\\ &= \frac{\alpha(|\Lambda|-1)(\alpha q-R(q))}{\left[R(q)+\alpha((|\Lambda|-2)q+\pi-p_{\bar{c}}+z)\right]^2} \leqslant 0,\end{aligned}$$

where we have used that $R(q) \geqslant \alpha q$, which holds as $\alpha \leqslant 1$ is the slope of the last piece-wise part of the function R, which extended up to the origin remains positive, in particular implying that the image of 0 (given by $R(q)-\alpha q$) is at least 0. We conclude $Q(z)$ is decreasing over $[p_{\bar{c}}, \pi]$, concluding that Equation (7.12) holds.

To finish the proof, suppose that such as $p_{\bar{c}}$ did not exist, as in Figure 7.16. Keep increasing the slope of the last pice-wise linear function until achieving the slop between q and the closest red dot placed at the left of q, namely $p_{\underline{c}}$, as in Figure 7.18 and set $\Lambda' := \Lambda \setminus \{\bar{c}\} \cup \{\underline{c}\}$.

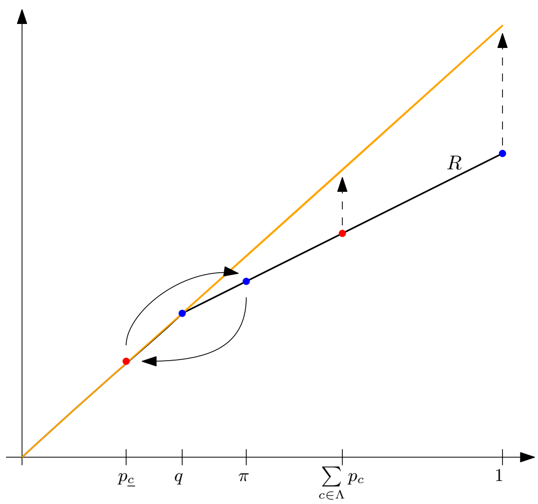


Fig. 7.18: Final construction Λ'

As in the previous cases, it can be proved that $F(p, R, \Lambda) \leqslant F(p, R, \Lambda')$. We omit the proof. □

Lemma 7.24. *Let $\pi \in [0,1]$ be fixed and $(p, R, \Lambda) \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$. Define*

$$\Gamma := \left\{k \in [C] \setminus \Lambda : p_k \leqslant \frac{1}{|\Lambda|}\sum_{c\in\Lambda} p_c\right\}.$$

There exists $(p', R') \in \Delta_\pi^C \times \mathcal{R}$ such that $F(p, R, \Lambda) \leqslant F(p', R', \Lambda \cup \Gamma)$.

Proof. Apply Lemmas 7.21 and 7.23 so the entry of p of value π is not included in Λ and for any $c \in \Lambda$, $p_c = q := \frac{1}{|\Lambda|}\sum_{c\in\Lambda} p_c$. In particular, the only values where F is increasing are $R(q)$ and $R(1)$. Define $\Gamma := p_c < q\}$. It follows that F is decreasing at $(R(p_c))_{c\in\Gamma}$. Replace R by

$$R(x) = \begin{cases} x\frac{R(q)}{q} & x \in [0, q] \\ \\ R(x) & x \in [q, \pi]. \end{cases}$$

Moreover, since $F(p, \alpha R, \Lambda) = F(p, R, \Lambda)$ for any $\alpha \neq 0$, redefine $R \equiv \frac{q}{R(q)}R$. Finally, since for any $q < p_c < 1$ the function F is decreasing on $R(p_c)$, replace R by,

$$R(x) = \begin{cases} x & x \in [0, q] \\ \\ q + (x - q) \cdot \frac{R(1)-q}{1-q} & x \in [q, \pi], \end{cases}$$

where we have used that $R(x) = x$ for any $x \leqslant q$ because of the previous scaling. The resulting function is illustrated in Figure 7.19 for $\Gamma = \{1, 2, 3\}$.

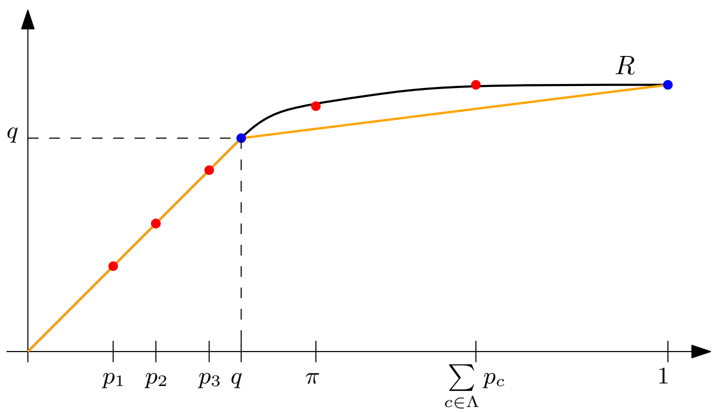


Fig. 7.19: Function R Lemma 7.24

Finally, we prove that $F(p, R, \Lambda) \leqslant F(p, R, \Lambda \cup \Gamma)$. For $z \in [0, q]$, consider

$$Q(z) := \frac{|\Lambda|R(q) + R(z)}{R(|\Lambda|q + z)} = \frac{|\Lambda|q + z}{q + (|\Lambda|q + z - q)\alpha},$$

where $\alpha = \frac{R(1)-q}{1-q}$. Remark the previous claim holds if and only if $Q(0) \leqslant Q(z)$, i.e., adding elements from Γ to Λ increases the part of F that depends on Λ. We obtain,

$$\frac{d}{dz}Q(z) = \frac{d}{dz}\left[\frac{|\Lambda|q + z}{q + (|\Lambda|q + z - q)\alpha}\right]$$

$$= \frac{(q + (|\Lambda|q + z - q)\alpha) - (|\Lambda|q + z)\alpha}{(q + (|\Lambda|q + z - q)\alpha)^2}$$
$$= \frac{q(1-\alpha)}{(q + (|\Lambda|q + z - q)\alpha)^2},$$

which is always positive. We conclude the proof. □

Lemma 7.25. *Let $\pi \in [0,1]$ be fixed and $(p, R, \Lambda) \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$. Then, there always exists $(p', R', \Lambda') \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$ such that,*

$$F(p', R', \Lambda') = \frac{\lambda}{\alpha(\lambda - 1) + 1} \cdot \frac{\alpha + (1-\alpha)q}{\alpha + (1-\alpha)Cq} =: \widehat{F}(q),$$

where $\lambda = |\Lambda'|$, $q = \frac{1}{\lambda}\sum_{c \in \Lambda'} p_c$, $\alpha = \frac{(R(1)-q)}{(1-q)}$, and $F(p, R, \Lambda) \leqslant F(p', R', \Lambda')$. In particular,

$$\arg\max_{q \in [0,1]} \widehat{F}(q) = \begin{cases} 0 \\ \frac{1-(C-\lambda)\pi}{\lambda} \end{cases}$$

Proof. Let $\pi \in [0,1]$ be fixed and $(p, R, \Lambda) \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$. Apply Lemmas 7.20, 7.21, 7.23 and 7.24 to construct $(p', R', \Lambda') \in \Delta_\pi^C \times \mathcal{R} \times 2^{[C]}$ such that,

$$F(p, R, \Lambda) \leqslant F(p', R', \Lambda')$$
$$\text{for any } c \in \Lambda', p'_c = q := \frac{1}{\lambda}\sum_{c \in \Lambda'} p_c$$
$$\text{for any } c \in [C] \setminus \Lambda, p'_c \geqslant q,$$
$$R'(x) = \begin{cases} x & x \in [0, q] \\ q + \alpha(x - q) & x \in [q, 1] \end{cases}$$

Moreover, $c^* \in [C]$ such that $p_{c^*} = \pi$, is not included in Λ'. It is not hard to see that,

$$F(p', R', \Lambda') = \widehat{F}(q) = \frac{\lambda}{\alpha(\lambda - 1) + 1} \cdot \frac{\alpha + (1-\alpha)q}{\alpha + (1-\alpha)Cq}.$$

Remark that $F(p', R', \Lambda')$ is decreasing on q. In particular, making $q \to q - \varepsilon$ increases the value of F. However, remark the value of q defines the kink of the function R'. In addition, any decreasing on q implies to decrease the value on the entries of p within Λ'. Since p is a probability distribution, the decrease of mass must be re-injected on all other entries whose values are below π, as we cannot modify the value of the highest-value entry of p. In conclusion, whenever solving

$$\max_{q\in[0,1]} \widehat{F}(q),$$

we obtain the solution

$$q = \begin{cases} 0 \\ \frac{1-(C-\lambda)\pi}{\lambda} \end{cases}$$

where the second case comes from attaining the constraint of maximizing all entries $c \in [C] \setminus \{\Lambda \cup \{c^*\}\}$ up to π. Remark that when the previous optimization problem we do not consider anymore the space of functions $\mathcal{R}$, in particular allowing for $q = 0$ to be a possible solution. $\square$

We are ready to prove Theorem 7.16.

Proof of Theorem 7.16. For $\lambda \in [C]$ and $\alpha \in [0,1]$, consider the function

$$\psi_\lambda(\alpha, q) = \frac{\lambda}{\alpha(\lambda-1)+1} \cdot \frac{\alpha + (1-\alpha)q}{\alpha + (1-\alpha)Cq}.$$

From Lemma 7.25, it follows,

$$\max_{p\in\Delta_\pi^C} \max_{R\in\mathcal{R}} \max_{\Lambda\subseteq[C]} \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \cdot \frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)} = \max_{\lambda\in[C]} \max_{\alpha\in[0,1]} \max_{q\in[0,1]} \psi_\lambda(\alpha, q). \tag{7.13}$$

We know that for $(\lambda, \alpha) \in [C] \times [0,1]$,

$$\arg\max_{q\in[0,1]} \psi_\lambda(\alpha, q) = \begin{cases} 0 \\ \frac{1-(C-\lambda)\pi}{\lambda} \end{cases}$$

Suppose $1 - (C-\lambda)\pi \leqslant 0$. It follows,

$$\max_{p\in\Delta_\pi^C} \max_{R\in\mathcal{R}} \max_{\Lambda\subseteq[C]} \frac{R(1)}{\sum_{c\in[C]} R(p_c)} \cdot \frac{\sum_{c\in\Lambda} R(p_c)}{R(\sum_{c\in\Lambda} p_c)} \leqslant \max_{\lambda\in[C]} \max_{\alpha\in[0,1]} \frac{\lambda}{\alpha(\lambda-1)+1} = \max_{\lambda\in[C]} \lambda.$$

Suppose $1 - (C-\lambda)\pi \geqslant 0$, i.e., $q \in [0, 1/C]$. We study the first order conditions of $\psi_\lambda(\alpha, q)$ over α. It follows,

$$\begin{aligned} \frac{d}{d\alpha}\psi_\lambda(\alpha, q) = -\frac{\lambda}{(\alpha\lambda-\alpha+1)^2(\alpha Cq-\alpha-Cq)^2} &[\alpha^2(C\lambda q^2 - C\lambda q - Cq^2 + Cq - \lambda q + \lambda + q - 1) \\ &+ \alpha(-2C\lambda q^2 + 2Cq^2 + 2\lambda q - 2q) + C\lambda q^2 - Cq^2 - Cq + q]. \end{aligned}$$

Imposing $\frac{d}{d\alpha}\psi_\lambda(\alpha, q) = 0$ we obtain the solutions,

$$\alpha_1 = \frac{2(-1+\lambda)q(-1+Cq) - \sqrt{4(-1+C)(-1+\lambda)q(-1+Cq)(-1+\lambda q)}}{2(q+Cq(-1+(-1+\lambda)q)},$$
$$\alpha_2 = \frac{2(-1+\lambda)q(-1+Cq) + \sqrt{4(-1+C)(-1+\lambda)q(-1+Cq)(-1+\lambda q)}}{2(q+Cq(-1+(-1+\lambda)q)}.$$

Since $q \leqslant 1/C$, it follows that $2(-1+\lambda)q(-1+Cq) \leqslant 0$ and, therefore, $\alpha_1 < 0$. The only possible solution being α_2, we show that either

1. $\alpha_2 \in [0,1]$ and then the stated value of $\psi_\lambda(q)$ for $q \in (0, \frac{\lambda-1}{\lambda(C-1)}]$ comes from plugging α_2 into Equation (7.13) or,

2. $\alpha_2 \geqslant 1$ and then Equation (7.13) is upper bounded by 1 for any $q \in (\frac{\lambda-1}{\lambda(C-1)}, \frac{1}{C}]$.

The first point is direct. For the second point, notice that $\alpha_2 \geqslant 1$ if and only if

$$\sqrt{4(C-1)(\lambda-1)q(Cq-1)(\lambda q-1)} \geqslant 2(\lambda-1)(p-1)(Cq-1) + 2(\lambda-1)q(1-Cq),$$

and, as $2(\lambda-1)(p-1)(Cq-1) + 2(\lambda-1)q(1-Cq) \geqslant 0$, this is equivalent to

$$\begin{aligned}
&4(C-1)(\lambda-1)q(Cq-1)(\lambda q-1) \geqslant (2(\lambda-1)(p-1)(Cq-1) + 2(\lambda-1)q(1-Cq))^2\\
\iff &4(\lambda-1)(1-q)(1-Cq)(\lambda((C-1)q-1)+1) \geqslant 0\\
\iff &q \in \left(\frac{\lambda-1}{\lambda(C-1)}, \frac{1}{C}\right].
\end{aligned}$$

Since $\alpha_2 \geqslant 1$, the optimal value of α is 1, yielding

$$\max_{\alpha\in[0,1]}\max_{q\in[0,1]} \psi_\lambda(\alpha, q) = \max_{q\in\left(\frac{\lambda-1}{\lambda(C-1)}, \frac{1}{C}\right]} \psi_\lambda(1, q) = 1.$$

To conclude the proof, set $\psi_\lambda(q) := \max_{\alpha\in[0,1]} \psi_\lambda(\alpha, q)$. The relaxed upper bound

$$\max_{\lambda\in[C]} \psi_\lambda\left(\frac{1-(C-\lambda)\pi}{C}\right) \leqslant C - \frac{1}{\pi},$$

is obtained through symbolic computation in Mathematica (with the Reduce function). Indeed, it can be verified that the inequality system

$$\frac{\lambda(Cq-1)}{\lambda q-1} - \frac{\lambda\left(Cq(\lambda q+\lambda-2) - 2\sqrt{(C-1)(\lambda-1)q(Cq-1)(\lambda q-1)} - 2\lambda q + q + 1\right)}{(C\lambda q-1)^2} \geqslant 0$$

for $q \in \left[0, \dfrac{\lambda-1}{C\lambda-\lambda}\right]$ and $\lambda \in [2, C-1]$, is always feasible. It follows that $\lambda(Cq-1)/(\lambda q-1) \geqslant \psi_\lambda(q)$ for $0 \leqslant q \leqslant (\lambda-1)/(C\lambda-\lambda)$. In particular, for $q=(1-(C-\lambda)\pi)/\lambda$, it follows that $C-1/\pi \geqslant \psi_\lambda(q)$ over $[1/(C-1), 1/(C-\lambda)]$. Similarly, since $C-1/\pi$ is greater than $\lambda=\psi_\lambda(q)$ whenever $\pi \geqslant 1/(C-\lambda)$, we conclude $C-1/\pi \geqslant \psi_\lambda(q)$ for any $\lambda \in [C]$ and $q \in [0,1]$. □

7.9.11 A technical lemma

The following technical lemma gives a sufficient condition for a polymatroid to have a price of opportunity fairness equal to 1. In particular, several of the posterior results in the stochastic setting will use it.

Lemma 7.26. *Let M be a polymatroid. Given a permutation $\sigma \in \Sigma([C])$, consider the sequence $\mathrm{r}(\sigma) = (\mathrm{r}_c(\sigma))_{c\in[C]}$ such that,*

$$\textit{for any } c \in [C], \mathrm{r}_c(\sigma) := \frac{\mathrm{r}(\sigma(1,\dots,c)) - \mathrm{r}(\sigma(1,\dots,c-1))}{\mathrm{r}(\sigma(c))},$$

where, $\mathrm{r}(\sigma(1,\dots,c))$ corresponds to the size of a maximum size allocation in the submatroid obtained by the groups in the first c entries of $\sigma([C])$. Whenever the sequences $\mathrm{r}(\sigma)$ for any $\sigma \in \Sigma([C])$, are all decreasing, it holds $\mathrm{PoF}(M) = 1$.

Proof. Let $\Lambda^* = \arg\max_{\Lambda\subseteq[C]} \frac{\sum_{c\in\Lambda}\mathrm{r}(c)}{\mathrm{r}(\Lambda)}$. We aim at proving that the monotonicity of the sequences $\{\mathrm{r}(\sigma), \sigma \in \Sigma([C])\}$ implies $\Lambda^* = [C]$, which yields $\mathrm{PoF}(M) = 1$. Without loss of generality, take $\sigma = \mathrm{I}_C$ to be the identity permutation (the same argument works for any other permutation). Denote

$$\rho_t := \frac{\mathrm{r}([t])}{\sum_{\ell\in[t]}\mathrm{r}(\ell)},$$

the competition index of the submatroid obtained by the the first t groups. Denoting $\mathrm{r}(0) = 0$, it follows,

$$\begin{aligned}
\rho_{t+1} - \rho_t &= \frac{\mathrm{r}([t+1])}{\sum_{\ell\in[t+1]}\mathrm{r}(\ell)} - \frac{\mathrm{r}([t])}{\sum_{\ell\in[t]}\mathrm{r}(\ell)} \\
&= \frac{\sum_{\ell\in[t+1]}\mathrm{r}(\ell) - \mathrm{r}(\ell-1)}{\sum_{\ell\in[t+1]}\mathrm{r}(\ell)} - \frac{\sum_{\ell\in[t]}\mathrm{r}(\ell) - \mathrm{r}(\ell-1)}{\sum_{\ell\in[t]}\mathrm{r}(\ell)} \\
&= \frac{\sum_{\ell\in[t]}[\mathrm{r}(t+1) - \mathrm{r}(t)]\mathrm{r}(\ell) - \mathrm{r}(t+1)[\mathrm{r}(\ell) - \mathrm{r}(\ell-1)]}{\left(\sum_{\ell\in[t+1]}\mathrm{r}(\ell)\right)\left(\sum_{\ell\in[t]}\mathrm{r}(\ell)\right)}.
\end{aligned}$$

Since $\mathrm{r}(\sigma)$ is decreasing, for any $s < t + 1$ it follows,

$$\frac{\mathrm{r}(s) - \mathrm{r}(s-1)}{\mathrm{r}(s)} \geqslant \frac{\mathrm{r}(t+1) - \mathrm{r}(t)}{\mathrm{r}(t+1)}.$$

In particular, $\mathrm{r}(t+1)[\mathrm{r}(s) - \mathrm{r}(s-1)] \geqslant [\mathrm{r}(t+1) - \mathrm{r}(t)]\mathrm{r}(s)$ and therefore, $\rho_t \geqslant \rho_{t+1}$. It follows that the optimal solution corresponds to $\Lambda^* = [C]$. □

7.9.12 Proof of Propositions 7.18 and 7.19

Proposition 7.18. Let $\omega = \omega(n)$ be a function such that $\omega(n) \to \infty$. Whenever $q \leqslant 1/(\omega n)$ or $q \geqslant \omega/n$, for any $p \in \Delta^C$, $\mathrm{PoF}(\mathrm{G}_{n,q}(p))$ converges to 1 with high probability as n grows.

Proof. We show this proposition by leveraging results from random graph theory. Suppose $q \leqslant 1/(\omega n)$. By Theorem 2.1 [FK16], $\mathrm{G}_{n,q}$ is a forest w.h.p.. It follows that, independent of the label realization, the maximal allocation contain all edges in the graph. Hence $\sum_{c\in[C]} \mathrm{r}(c) = \mathrm{r}([C])$, which implies the matroid has PoF equal to 1 as the independence index is equal to 1.

Suppose $q \geqslant \omega/n$. For any $c \in [C]$ the subgraph induced by considering only the subset E_c over $\mathrm{G}_{n,q}$ is distributed according to $\mathrm{G}_{n,p_c q}$. Since $p_c q \geqslant p_c\omega/n$, with $p_c\omega \to \infty$ arbitrarily slow, Theorem 2.14 [FK16] states that w.h.p. $\mathrm{G}_{n,p_c q}$ has a giant component of size $(1 - \frac{x}{p_c\omega})n$, for a fixed $x \in [0, 1]$. In particular, $\mathrm{r}(c) = (1 - \frac{x}{p_c\omega})n$ as connected components contain spanning trees. Intersecting the events over all $c \in [C]$ we obtain, w.h.p. as n goes to infinity,

$$\rho(\mathrm{G}_{n,q}(p)) = \frac{\mathrm{r}([C])}{\sum_{c\in[C]} \mathrm{r}(c)} = \frac{n}{\sum_{c\in[C]} n} + o(1) = \frac{1}{C} + o(1),$$

which from Proposition 7.14 shows that PoF is also equal to 1 w.h.p.. □

Proposition 7.19. Let $\omega = \omega(n)$ be a function such that $\omega(n) \to \infty$ arbitrarily slow as $n \to \infty$. Whenever $q \leqslant 1/(\omega n^{3/2})$ or $q \geqslant \omega \log(n)/n$, for any $p \in \Delta^C$, $\mathrm{PoF}(\mathrm{B}_{n,\beta,q}(p))$ converges to 1 with high probability as n grows.

Proof. Suppose $q \geqslant \omega \log(n)/n$. Let $\Lambda \subseteq [C]$, we have that $\sum_{c\in[C]} |E_c|$ is a sum of independent bernouli random variables (the E_c are disjoints), hence it has an expected value of $n \sum_{c\in\Lambda} p_c$ and Hoeffding's concentration inequality show that

$$\mathbb{P}\Big(|\sum_{c\in\Lambda} |E_c| - n \sum_{c\in\Lambda} p_c| > \sqrt{n\log(n)}\Big) \leqslant 2\exp\Big(-2\frac{n\log(n)}{n}\Big) = \frac{2}{n^2} \xrightarrow[n\to\infty]{} 0.$$

Since $q \geqslant \omega \log(n)/n$, Theorem 6.1 [FK16] states that w.h.p. for any $\Lambda \subseteq [C]$, the subgraph considering only the vertices in Λ on the left-hand side has a matching of size $\min\{\beta n, \sum_{c\in\Lambda} p_c n\}$, therefore, $\mathrm{r}(\Lambda) = \min\{\beta n, \sum_{c\in\Lambda} p_c n\}$. We will conclude by applying Lemma 7.26. As usual, w.l.o.g. consider $\sigma = \mathrm{I}_L$. For any $c \in [C-1]$, it follows,

$$\mathrm{r}_{c+1}(\sigma) = \frac{\min\{\beta n, \sum_{c'\in[c+1]} p_{c'} n\} - \min\{\beta n, \sum_{c'\in[c]} p_{c'} n\}}{\min\{\beta n, p_c n\}}.$$

In particular, as $\sum_{c'\in[c]} p_{c'}$ is increasing in c, the sequence $\mathrm{r}_{c+1}(\sigma)$ initially consists on only 1 (given by all times that $\mathrm{r}_{c+1}(\sigma) \leqslant \beta$), eventually some value between 0 and 1 (given by the first time that $\mathrm{r}_{c+1}(\sigma) \geqslant \beta \geqslant \mathrm{r}_c(\sigma)$), and finally a sequence of only zeros (given by all times when $\mathrm{r}_{c+1}(\sigma) > \beta$). In particular, the sequence is decreasing, concluding the proof. □

Remark. In both graphic and transversal random matroids, taking the same $q \in [0,1]$ for all colors is done without loss of generality. Indeed, a coupling argument based on stochastic dominance allows us to consider edge probabilities q_c per color c and to obtain the same results.

7.10 Price of Fairness under Other Fairness Notions

7.10.1 Weighted Fairness

The main fairness definition that we have used is opportunity fairness. We now discuss how this specific fairness notion relates with other fairness concepts, in particular with maximin fairness and proportionality in [BFT11] and equitability in [Car+12]. We can think more generally about group fairness in terms of what amount of social welfare protected group of agents are entitled to. Should each group be entitled to the same amount as others, or proportionally to their size? We introduce weighted fairness, where the weights correspond to group entitlement:

Definition 7.27. Let $(w_c)_{c\in[C]} \in \mathbb{R}_+^C$ be a fixed weights vector. An allocation $x \in \mathbb{R}_+^C$ is w-fair if for any $i, j \in [C]$, $x_i/w_i = x_j/w_j$.

As an example, Figure 7.20 illustrates for a 2-colored matroid three fairness notions mentioned in the paper that are now framed as specific instances of weighted fairness.

1. **Equitability**

   $w_c = 1$ for $c \in [C]$,

2. **Proportional fairness**

   $w_c = |E_c|$ for $c \in [C]$,

3. **Opportunity fairness**

   $w_c = \mathrm{r}(c)$ for $c \in [C]$.

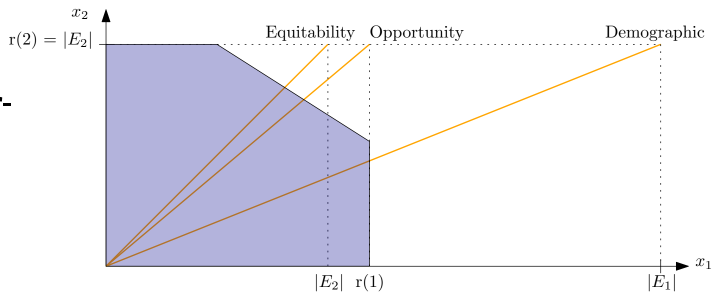


Fig. 7.20: Weighted Fairness for matroid with two groups

Compared to other weighted fairness notions, opportunity fairness remains bounded because the weights depend on the structure of the polymatroid M, while the weights of proportional fairness and equitability are independent of M and arbitrarily bad examples can easily be constructed.

Proposition 7.28. *The price of proportional fairness and the price of equitability are unbounded in the worst-case, even when allowing for fractional allocations.*

Proof. Take a ground set with n agents of color 1 and n agents of color 2, and consider the feasible allocations where at most $n/2$ resources can be allocated to individuals of color 2 but only one resource may be allocated to individuals of color 1. This is a partition matroid, where the partition corresponds exactly to the color partition. Now, the optimal proportionally fair allocation, as well as the optimal equitable allocation, is $x_1 = 1$ and $x_2 = 1$, for a total of 2 resources allocated. Because the optimal allocation is of size $n + 1$, the price of fairness is $(n + 1)/2$, which goes to ∞ as $n \to \infty$. The price of fairness in both cases is unbounded. $\square$

Another common concept of fairness to divide resources, used in transferable utility cooperative game theory, is that of Shapley value [OR95]: it is the unique utility transfer that satisfies axioms of symmetry, additivity, nullity and efficiency. It can be shown that for $\Sigma([C])$ the set of permutations over $[C]$, the Shapley value of group c is

$$\varphi_c := \frac{1}{C!}\sum\nolimits_{\sigma\in\Sigma([C])} \mathrm{r}(\{i \in [C] \mid \sigma(i) < \sigma(c)\} \cup \{c\}) - \mathrm{r}(\{i \in [C] \mid \sigma(i) < \sigma(c)\}),$$

that is to say φ_c is the expected marginal contribution of group c when groups are prioritized according to σ a uniformly drawn random permutation. When $w_c = \varphi_c$, we say that an allocation is Shapley fair.

The allocation problem we study can be seen as a type of non transferrable utility game, and as such there is no reason in general for the allocation $(\varphi_1, \dots, \varphi_C)$ to be realizable. Nonetheless, from the polymatroid characterization of M do have this property:

Proposition 7.29. *The allocation $(\varphi_1, \dots, \varphi_C)$ is always feasible.*

Proof. For a given permutation σ, the marginal contribution allocation x^σ wher $x^\sigma_c = \mathrm{r}(\{i \in [C] \mid \sigma(i) < \sigma(c)\} \cup \{c\}) - \mathrm{r}(\{i \in [C] \mid \sigma(i) < \sigma(c)\})$ is always feasible. Hence, the Shapley allocation $(\varphi_1, \dots, \varphi_C)$ the barycenter of all the x^σ, which belong to the Pareto front by definition. Moreover, by the polymatroid characterization, the Pareto front is convex [HH02]. Hence the barycenter, being a convex combination, is also feasible. We note that the x^σ are the extreme points of the Pareto front. $\square$

From the efficiency of the Shapley allocation, it is immediate that the Shapley Price of Fairness is always 1 for colored matroids.

In the semi-random model of Theorem 7.16, we can easily show the following property:

Proposition 7.30. *For any distribution $p \in \Delta^C$ of the agents colors, in the large market setting with $\liminf \mathrm{r}_n([C]) = \Omega(n)$, we have that the price of proportional fairness converges to 1 with high probability.*

Proof. Let S_n be any maximal allocation, taken independently of the random coloring. We have that $|S_n| = \Omega(n)$ and $|S_n \cap E_c|$ concentrates towards $p_c|S_n|$ by Hoeffding's inequality. We also have that $|E_c|$ concentrates around $p_c n$. Hence $|S_n \cap E_c|/|E_c|$ concentrates around $|S_n|/n$, which is independent of the colors, and therefore is a proportionally fair allocation. Moreover S_n is maximal by definition, so the price of fairness is equal to 1. $\square$

This shows that the price of proportional fairness goes from $\mathrm{PoF} = +\infty$ in the adversarial setting to $\mathrm{PoF} = 1$ in the semi-random setting.

Finally let us mention another fairness definition.

7.10.2 Leximin Fairness

Requiring that a fair allocation satisfies exactly $x_i/w_i = x_j/w_j$ can be considered wasteful, as it is possible to improve the total social welfare without making any group worst off. Maximin fairness [BFT11], also called egalitarian rule or Rawlsian fairness, corresponds to ensuring that the worst off group has the best allocation possible. In other words, an allocation is maximin fair if it maximizes $\min_{x\in M} x_c$, or with entitlement w, maximizes $\min_{x\in M} x_c/w_c$. Most of the time there are multiple maximin fair feasible allocations, and thus one may seek to maximize the second minimum, and so forth. This is called the leximin rule, and has also been studied in the social choice literature [DG77; DG78].

For a vector $x = (x_1, \ldots, x_C)$, we denote the ordered coordinates by $x_{(1)} \geqslant x_{(2)} \geqslant \ldots \geqslant x_{(C)}$. We say that a vector $x = (x_1, \ldots, x_C)$ is leximin larger than $y = (y_1, \ldots, y_C)$ if $x_{(C)} \geqslant y_{(C)}$, or $x_{(C)} = y_{(C)}$ and $x_{C-1} \geqslant y_{(C-1)}$, or $x_{(C)} = y_{(C)}$ and $x_{C-1} = y_{(C-1)}$ and $x_{C-2} = y_{(C-2)}$ and so on. The leximin order is a total preorder. Leveraging the more general notion of weighted fairness, we have the following definition:

Definition 7.31. For a weight vector $w \in \mathbb{R}_+^C$, an allocation $(x_1, \ldots, x_c)$ is said to be w-lexmaxmin fair if $\left(\frac{x_1}{w_1}, \ldots, \frac{x_C}{w_C}\right)$ is maximal according to the leximin order for $x \in M$.

Clearly, the w-lexmaxmin fair allocation is w-maxmin fair. It is also Pareto efficient, and therefore by the polymatroid characterization Proposition 7.7 achieves maximal social welfare: the price of w-lexmaxmin fairness is always 1 in C-colored matroids.

8 Conclusion

This thesis has examined the impact of fairness constraints on five standard resource allocation problems, emphasizing how equity considerations modify both algorithm design and performance guarantees. From online advertising to refugee resettlement, the problems explored span a wide range of practical domains and mathematical settings, ranging from sequential and statistical models to strategic and combinatorial environments.

A recurring theme that is at the heart of all chapters is the trade-off between efficiency and fairness. While it is clear that, in worst-case scenarios, fairness can significantly reduce achievable utility or social welfare, we show that this loss can often be substantially mitigated through careful algorithmic design. Moreover, we have demonstrated that in many realistic settings (e.g. with stochastic structure Chapter 7), fairness comes at a modest or even negligible cost, and sometimes even for free (Chapter 6). Achieving equity in decision-making is undoubtedly a challenge, but this thesis shows that it is rarely an insurmountable one, especially when decision-makers are both capable and willing to adapt their methods to the specific structure of the problem at hand. That said, in some dynamic or biased environments, fairness constraints can also improve long-term utility by encouraging exploration or correcting implicit biases in the objective to optimize itself (see, e.g., [KMR18]).

Future Works

Going forward, I am especially interested in further developing the connections between fairness and competitive analysis, particularly within the framework of prophet inequalities. While fairness has now been widely explored in bandit and learning-based settings, its role in online algorithms from a worst-case perspective remains underexplored. Only a few recent works have addressed fairness in the context of prophet inequalities or online matching more broadly [MX20; AK22; Cor+21]. In Section 5.2 we suggest a promising direction by incorporating fairness penalties into the objective function, which is an ongoing work. An alternative approach would be to enforce fairness through hard constraints, for example, requiring that selected elements satisfy group-level quotas. While matroid constraints have been widely studied in the prophet literature [KW12b], covering constraints remain far less understood and could open up a new line of research.

Another direction – perhaps even more critical in the long run – is understanding the strategic response of agents to fairness-aware systems. As fairness constraints become more prominent in policy and platform design, rational agents will begin to react strategically to these rules. This could involve manipulating input information, misreporting group membership, or gaming the system to exploit fairness-induced behavior. These concerns have been studied in recent work in strategic classification [Har+16a], where fair classifiers are shown to induce behavioral shifts that may undermine fairness in the long term [Liu+18; HIL19]. Similar incentives could emerge in allocation problems: for example, in repeated auctions with group-fairness constraints, how will bidders adapt? Could they deliberately inflate prices for underrepresented groups to deplete their adversaries' budgets? Exploring these interactions between fairness and strategic behavior, especially in repeated or dynamic settings, could be crucial to predict future agents' behavior.

Closing Reflections

There are two broad positive takeaways from this thesis: one related to applied mathematics research, the other societal.

While they for sure deserve attention due to their ethical grounding, fairness constraints, at the end of the day, are just yet another constraints. So why is that they deserve such an important treatment from a theory perspective? Theory problems in applied mathematics often draw inspiration from real-world problems. Much like the way that the problem of kidney exchange inspired foundational results in random graph theory, fairness can serve as a lens through which we motivate and isolate meaningful mathematical questions in the space of all possible resource allocation questions. As an example, let us quote a sentence from [BFT11]: "[...] when the set of achievable 'utilities' is a polymatroid, all Pareto resource allocations are efficient. This is, unfortunately, a somewhat restrictive condition, and a general class of resource allocation problems that satisfy this condition is not known.". In Chapter 7 we show how a polymatroid structure naturally emerges when considering group-fairness, which could be a trigger for renewed interest in studying polymatroids. This suggests that fairness considerations may guide us toward tractable and well-structured subclasses of otherwise complex allocation problems.

From an applied perspective, why should anyone be interested in theoretical results on fair allocations? It is important to remember that this thesis is first and foremost a mathematical study, and there are no prescriptive claims about what fairness should be, or how public policy ought to enforce it. Rather, we advance the understanding of how fairness requirements limit efficiency in specific settings, and characterize what exactly is achievable in fair resource allocation. In some sense, we are trying to delimit what the realm of the possible is, which serves as a useful foundation for policymakers: any normative decision about how fairness should be enforced must rest on what is actually achievable. If this thesis helps clarify those boundaries, even if we only paint a small corner, it may help public institutions in shaping more realistic, enforceable, and ultimately effective policies.

However, it would be naïve to assume that the relationship between fairness and applied mathematics research is purely positive. These domains have distinct objectives: theoretical research aims for elegant, mathematically challenging, generalizable abstractions, while fairness stems from concrete experiences of injustice and marginalization. As such, formal models inevitably reflect a bias toward fairness notions that align with tractable objectives or well-understood mathematical structures. These may appeal to researchers, but not necessarily to those most affected by allocation decisions. In that sense, research on fair allocation may itself be unfair, by prioritizing what is analytically convenient over what is socially urgent.

Moreover, there is a risk that formal fairness definitions may be appropriated in ways that aggravate or sustain harm. For example, some fairness notions, such as predictive parity [Cho17], align closely with utility maximization and are more easily satisfied by standard classifiers. But satisfying such a criterion does not guarantee desired fairness in specific applications, and may even mask underlying inequities. Because different fairness definitions are often mutually incompatible, it becomes possible to selectively adopt the one that best fits an decision-maker objectives, without meaningfully altering outcomes. In extreme cases, one could imagine decision-makers optimizing solely for utility, and then retrospectively searching for the fairness definition under which their decisions appear least discriminatory, providing a performative fairness certificate. This would turn fairness into an ex-post rationalization of decisions, rather than a guiding principle.

To avoid this kind of distortion, fairness research must remain grounded in the reality of discrimination suffered by individuals. This requires deeper engagement with the social sciences and humanities, collaboration with civil society organizations, and dialogue with public institutions. Theoretical work has much to contribute, but it must remain at the service of the values and priorities of the communities it hopes to serve, and not lose itself in a self-satisfying pursuit of elegance, serving nothing beyond its own reflection.

Bibliography

[APS11] Yasin Abbasi-yadkori, Dávid Pál, and Csaba Szepesvári. "Improved Algorithms for Linear Stochastic Bandits". In: *Advances in Neural Information Processing Systems*. 2011 (cit. on pp. 30, 75).

[AD16a] Shipra Agrawal and Nikhil Devanur. "Linear Contextual Bandits with Knapsacks". In: *Advances in Neural Information Processing Systems*. 2016 (cit. on p. 27).

[AD19] Shipra Agrawal and Nikhil R. Devanur. "Bandits with Global Convex Constraints and Objective". In: *Operation Research* (2019) (cit. on pp. 95, 96, 99).

[AD15] Shipra Agrawal and Nikhil R. Devanur. "Fast Algorithms for Online Stochastic Convex Programming". In: *Proceedings of the 2015 Annual ACM-SIAM Symposium on Discrete Algorithms*. 2015 (cit. on pp. 25, 27).

[AD16b] Shipra Agrawal and Nikhil R. Devanur. "Linear Contextual Bandits with Knapsacks". In: *Neural Information Processing Systems*. 2016 (cit. on p. 95).

[AG12] Shipra Agrawal and Navin Goyal. "Analysis of Thompson Sampling for the Multi-armed Bandit Problem". In: *Conference on Learning Theory*. 2012 (cit. on p. 98).

[Aha+21] Narges Ahani, Paul Gölz, Ariel D. Procaccia, Alexander Teytelboym, and Andrew C. Trapp. "Dynamic Placement in Refugee Resettlement". In: *Proceedings of the 22nd ACM Conference on Economics and Computation* (2021) (cit. on pp. 25, 42, 257).

[Akb+20] Mohammad Akbarpour, Julien Combe, Yinghua He, et al. "Unpaired kidney exchange: Overcoming double coincidence of wants without money". In: *Proceedings of the 21st ACM Conference on Economics and Computation*. 2020, pp. 465–466 (cit. on p. 257).

[ADK24] Mohammad Akbarpour, Piotr Dworczak, and Scott Duke Kominers. "Redistributive Allocation Mechanisms". In: *Journal of Political Economy* 132.6 (2024), pp. 1831–1875 (cit. on p. 204).

[Akb+22] Mohammad Akbarpour, Adam Kapor, Christopher Neilson, Winnie Van Dijk, and Seth Zimmerman. "Centralized school choice with unequal outside options". In: *Journal of Public Economics* 210 (2022), p. 104644 (cit. on p. 257).

[AN19] Mohammad Akbarpour and Afshin Nikzad. "Approximate Random Allocation Mechanisms". In: *Microeconomics: Welfare Economics & Collective Decision-Making eJournal* (2019) (cit. on p. 263).

[Ala11] Saeed Alaei. “Bayesian Combinatorial Auctions: Expanding Single Buyer Mechanisms to Many Buyers”. In: *IEEE Annual Symposium on Foundations of Computer Science* (2011) (cit. on pp. 141, 195).

[Ala+22] Saeed Alaei, Ali Makhdoumi, Azarakhsh Malekian, and Rad Niazadeh. “Descending Price Auctions with Bounded Number of Price Levels and Batched Prophet Inequality”. In: *ACM Conference on Economics and Computation* (2022) (cit. on p. 196).

[Ali+19a] Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, et al. “Discrimination through Optimization: How Facebook’s Ad Delivery can Lead to Biased Outcomes”. In: *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. 2019 (cit. on p. 25).

[Ali+19b] Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, et al. “Discrimination through optimization: How Facebook’s Ad delivery can lead to biased outcomes”. In: *Proceedings of the ACM on human-computer interaction* 3.CSCW (2019), pp. 1–30 (cit. on p. 257).

[Als08] Brian Alspach. “The wonderful Walecki construction”. In: *Bull. Inst. Combin. Appl* 52.52 (2008), pp. 7–20 (cit. on p. 281).

[ÁM07] Francisco Álvarez and Cristina Mazón. “Comparing the Spanish and the discriminatory auction formats: A discrete model with private information”. In: *European Journal of Operational Research* 179.1 (2007), pp. 253–266 (cit. on p. 227).

[AAT19] Sanae Amani, Mahnoosh Alizadeh, and Christos Thrampoulidis. “Linear Stochastic Bandits Under Safety Constraints”. In: *Neural Information Processing Systems*. 2019 (cit. on p. 95).

[AVV07] Glen Douglas Anderson, Mavina Vamanamurthy, and Matti Vuorinen. “Generalized convexity and inequalities”. In: *Journal of Mathematical Analysis and Applications* 335 (2007), pp. 1294–1308 (cit. on p. 224).

[Ari99] Aristotle. *Nicomachean Ethics*. Ed. and trans. by Terence Irwin. 2nd. Original work written ca. 350 B.C.E. Indianapolis: Hackett Publishing Company, 1999 (cit. on p. 3).

[ABN08] Barry C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Classics in Applied Mathematics A First Course in Order Statistics*. Society for Industrial and Applied Mathematics, 2008 (cit. on p. 248).

[AM21] Nick Arnosti and Will Ma. “Tight Guarantees for Static Threshold Policies in the Prophet Secretary Problem”. In: *Conference on Economics and Computation* (2021) (cit. on pp. 142, 160, 161, 189).

[AK22] Makis Arsenis and Robert D. Kleinberg. “Individual Fairness in Prophet Inequalities”. In: *Proceedings of the 23rd ACM Conference on Economics and Computation* (2022) (cit. on p. 311).

[AG79] Claude d'Aspremont and Louis-André Gérard-Varet. "Incentives and incomplete information". In: *Journal of Public Economics* 11.1 (1979), pp. 25–45 (cit. on p. 205).

[ACL13] Susan Athey, Dominic Coey, and Jonathan Levin. "Set-Asides and Subsidies in Auctions". In: *American Economic Journal: Microeconomics* 5.1 (2013) (cit. on p. 204).

[Atk70] Anthony B. Atkinson. "On the measurement of inequality". In: *Journal of Economic Theory* 2.3 (1970), pp. 244–263 (cit. on p. 203).

[ACF02] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. "Finite-time analysis of the multiarmed bandit problem". In: *Machine learning* (2002) (cit. on p. 98).

[Aum60] Robert J. Aumann. "Acceptable Points in General Cooperative n-Person Games". In: *Contributions to the Theory of Games IV*. Ed. by A. W. Tucker and R. D. Luce. Vol. 40. Annals of Mathematics Studies. Princeton University Press, 1960, pp. 287–324 (cit. on p. 14).

[Aus+14] Lawrence M. Ausubel, Peter Cramton, Marek Pycia, Marzena Rostek, and Marek Weretka. "Demand Reduction and Inefficiency in Multi-Unit Auctions". In: *Review of Economic Studies* 81 (2014), pp. 1366–1400 (cit. on pp. 201, 225).

[AKW14] Pablo D. Azar, Robert Kleinberg, and S. Matthew Weinberg. "Prophet inequalities with limited information". In: *ACM-SIAM Symposium on Discrete Algorithms*. 2014 (cit. on p. 196).

[Azi20] Haris Aziz. "Simultaneously Achieving Ex-ante and Ex-post Fairness". In: *ArXiv* abs/2004.02554 (2020) (cit. on p. 263).

[BF22] Moshe Babaioff and Uriel Feige. "Fair Shares: Feasibility, Domination and Incentives". In: *Proceedings of the 23rd ACM Conference on Economics and Computation* (2022) (cit. on p. 10).

[BKS18a] Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. "Bandits with Knapsacks". In: *J. ACM* (2018) (cit. on p. 27).

[BKS18b] Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. "Bandits with Knapsacks". In: *Journal of the ACM* (2018) (cit. on p. 95).

[BB05] Mark Bagnoli and Ted Bergstrom. "Log-Concave Probability and Its Applications". In: *Economic Theory* 26.2 (2005), pp. 445–469 (cit. on pp. 218, 231, 250, 251).

[BLM20] Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. "Dual Mirror Descent for Online Allocation Problems". In: *Proceedings of the 37th International Conference on Machine Learning*. 2020 (cit. on p. 25).

[BLM21] Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. "Regularized Online Allocation Problems: Fairness and Beyond". In: *Proceedings of the 38th International Conference on Machine Learning*. 2021 (cit. on pp. 25–27, 29, 34–36, 40, 41, 45, 54).

[BBW14] Santiago R. Balseiro, Omar Besbes, and Gabriel Y. Weintraub. “Repeated Auctions with Budgets in Ad Exchanges: Approximations and Design”. In: *Columbia Business School Research Paper Series* (2014) (cit. on p. 45).

[Ban+23] Sayan Bandyapadhyay, Fedor V. Fomin, Tanmay Inamdar, and Kirill Simonov. “Proportionally Fair Matching with Multiple Groups”. In: *Graph-Theoretic Concepts in Computer Science: 49th International Workshop*. 2023, pp. 1–15 (cit. on pp. 257, 273).

[BHN19] Solon Barocas, Moritz Hardt, and Arvind Narayanan. *Fairness and Machine Learning*. fairmlbook.org, 2019 (cit. on p. 25).

[BHN23] Solon Barocas, Moritz Hardt, and Arvind Narayanan. *Fairness and Machine Learning: Limitations and Opportunities*. MIT Press, 2023 (cit. on pp. 18, 19, 261).

[BSH23] Dorian Baudry, Kazuya Suzuki, and Junya Honda. “A General Recipe for the Analysis of Randomized Multi-Armed Bandit Algorithms”. In: *ArXiv : 2303.06058* (2023) (cit. on p. 98).

[BG86] Dimitri P. Bertsekas and Robert G. Gallager. “Data Networks”. In: 1986 (cit. on p. 16).

[BFT11] Dimitris Bertsimas, Vivek F. Farias, and Nikolaos Trichakis. “The Price of Fairness”. In: *Oper. Res.* 59 (2011), pp. 17–31 (cit. on pp. 258, 272, 307, 310, 312).

[Bey+11] Alina Beygelzimer, John Langford, Lihong Li, Lev Reyzin, and Robert Schapire. “Contextual Bandit Algorithms with Supervised Learning Guarantees”. In: *International Conference on Artificial Intelligence and Statistics*. 2011 (cit. on pp. 102, 111).

[BR91] Sushil Bikhchandani and John G. Riley. “Equilibria in open common value auctions”. In: *Journal of Economic Theory* 53.1 (1991), pp. 101–130 (cit. on p. 202).

[BM01a] Anna Bogomolnaia and Hervé Moulin. “A New Solution to the Random Assignment Problem”. In: *Journal of Economic Theory* 100.2 (2001), pp. 295–328 (cit. on pp. 1, 11).

[BM01b] Anna Bogomolnaia and Hervé Moulin. “A new solution to the random assignment problem”. In: *Journal of Economic Theory* 100.2 (2001), pp. 295–328 (cit. on p. 263).

[BE98] Allan Borodin and Ran El-Yaniv. *Online computation and competitive analysis*. Cambridge University Press, 1998 (cit. on p. 137).

[Bra+16] F. Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. “Handbook of Computational Social Choice”. In: 2016 (cit. on p. 9).

[BPV24] Johannes Brustle, Sebastian Perez-Salazar, and Victor Verdugo. *Splitting Guarantees for Prophet Inequalities via Nonlinear Systems*. 2024. arXiv: 2406.17767 (cit. on p. 196).

[BCL13] Sébastien Bubeck, Nicolo Cesa-Bianchi, and Gábor Lugosi. "Bandits with heavy tail". In: *IEEE Transactions on Information Theory* (2013) (cit. on p. 110).

[Bub+18] Sébastien Bubeck, Michael B. Cohen, James R. Lee, and Yin Tat Lee. "Metrical task systems on trees via mirror descent and unfair gluing". In: *ACM-SIAM Symposium on Discrete Algorithms*. 2018 (cit. on p. 137).

[BC22] Archit Bubna and Ashish Chiplunkar. "Prophet Inequality: Order selection beats random order". In: *ACM Conference on Economics and Computation* (2022) (cit. on p. 195).

[BJS10] Niv Buchbinder, Kamal Kumar Jain, and Mohit Singh. "Secretary Problems via Linear Programming". In: *Mathematics of Operations Research*. 2010 (cit. on p. 196).

[Büc+22] Michael Bücker, Gero Szepannek, Alicja Gosiewska, and Przemyslaw Biecek and. "Transparency, auditability, and explainability of machine learning models in credit scoring". In: *Journal of the Operational Research Society* 73.1 (2022), pp. 70–90 (cit. on p. 1).

[Bud+13] Eric Budish, Yeon-Koo Che, Fuhito Kojima, and Paul R. Milgrom. "Designing Random Allocation Mechanisms: Theory and Applications". In: *The American Economic Review* 103 (2013), pp. 585–623 (cit. on p. 263).

[BK96] Jeremy Bulow and Paul Klemperer. "Auctions Versus Negotiations". In: *The American Economic Review* 86.1 (1996), pp. 180–194 (cit. on p. 225).

[BK02] Jeremy Bulow and Paul Klemperer. "Prices and the Winner's Curse". In: *The RAND Journal of Economics* 33.1 (2002), pp. 1–21 (cit. on p. 202).

[BG18a] Joy Buolamwini and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification". In: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. 2018 (cit. on pp. 1, 25).

[BG18b] Joy Buolamwini and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification". In: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. Vol. 81. Proceedings of Machine Learning Research. PMLR, 23–24 Feb 2018, pp. 77–91 (cit. on pp. 257, 265).

[Bur+21] Archana Bura, Aria HasanzadeZonuzy, Dileep M. Kalathil, Srinivas Shakkottai, and Jean-François Chamberland. "DOPE: Doubly Optimistic and Pessimistic Exploration for Safe Reinforcement Learning". In: *Neural Information Processing Systems*. 2021 (cit. on p. 96).

[Bur17] Robin Burke. "Multisided Fairness for Recommendation". In: *In Proceedings of the Workshop on Fairness, Accountability and Trans- parency in Machine Learning*. Available as arXiv:1707.00093. 2017. arXiv: `1707.00093 [cs.CY]` (cit. on p. 25).

[CV10] Toon Calders and Sicco Verwer. “Three naive Bayes approaches for discrimination-free classification”. In: *Data Mining and Knowledge Discovery* 21.2 (2010), pp. 277–292 (cit. on p. 19).

[Căl+11] Gruia Călinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. “Maximizing a Monotone Submodular Function Subject to a Matroid Constraint”. In: *SIAM J. Comput.* (2011), pp. 1740–1766 (cit. on p. 288).

[Cap+13] Olivier Cappé, Aurélien Garivier, Odalric-Ambrym Maillard, Rémi Munos, and Gilles Stoltz. “Kullback-Leibler Upper Confidence Bounds for Optimal Sequential Allocation”. In: *Annals of Statistics* (2013) (cit. on p. 98).

[Car+12] Ioannis Caragiannis, Christos Kaklamanis, P. Kanellopoulos, and Maria Kyropoulou. “The Efficiency of Fair Division”. In: *Theory of Computing Systems* (2012), pp. 589–610 (cit. on pp. 258, 272, 307).

[Car+16] Ioannis Caragiannis, David Kurokawa, Hervé Moulin, et al. “The Unreasonable Fairness of Maximum Nash Welfare”. In: *Proceedings of the 2016 ACM Conference on Economics and Computation* (2016) (cit. on p. 11).

[Car+19] Ioannis Caragiannis, David Kurokawa, Hervé Moulin, et al. “The unreasonable fairness of maximum Nash welfare”. In: 2019 (cit. on p. 272).

[Car+23] Emil Carlsson, D. Basu, Fredrik D. Johansson, and Devdatt P. Dubhashi. “Pure Exploration in Bandits with Linear Constraints”. In: *ArXiv : 2306.12774* (2023) (cit. on p. 96).

[CLP22] Rémi Castera, Patrick Loiseau, and Bary S. R. Pradelski. “Statistical Discrimination in Stable Matchings”. In: 2022 (cit. on p. 273).

[CCK23] Matteo Castiglioni, A. Celli, and Christian Kroer. “Online Learning under Budget and ROI Constraints and Applications to Bidding in Non-Truthful Auctions”. In: *ArXiv : 2302.01203* (2023) (cit. on p. 96).

[CCK22] Matteo Castiglioni, Andrea Celli, and Christian Kroer. “Online Learning with Knapsacks: the Best of Both Worlds”. In: *Proceedings of the 39th International Conference on Machine Learning*. 2022 (cit. on p. 27).

[Cas+22a] Matteo Castiglioni, Andrea Celli, Alberto Marchesi, Giulia Romano, and Nicola Gatti. *A Unifying Framework for Online Optimization with Long-Term Constraints*. Available as Arxiv:2209.07454. 2022 (cit. on p. 27).

[Cas+22b] Matteo Castiglioni, Andrea Celli, Alberto Marchesi, Giulia Romano, and Nicola Gatti. “A Unifying Framework for Online Optimization with Long-Term Constraints”. In: *Advances in Neural Information Processing Systems*. 2022 (cit. on p. 96).

[Cel+21] L. Elisa Celis, Lingxiao Huang, Vijay Keswani, and Nisheeth K. Vishnoi. “Fair Classification with Noisy Protected Attributes: A Framework with Provable Guarantees”. In: *Proceedings of the 38th International Conference on Machine Learning*. 2021 (cit. on p. 27).

[Cel+22] Andrea Celli, Riccardo Colini-Baldeschi, Christian Kroer, and Eric Sodomka. “The Parity Ray Regularizer for Pacing in Auction Markets”. In: *Proceedings of the ACM Web Conference*. Virtual Event, Lyon, France, 2022 (cit. on pp. 25–27, 29, 30, 34–36, 38, 40, 45, 53, 54).

[CC20] Tommaso Cesari and Roberto Colomboni. “A Nearest Neighbor Characterization of Lebesgue Points in Metric Measure Spaces”. In: *Mathematical Statistics and Learning* 3 (2020) (cit. on p. 49).

[CS21] H. Chang and R. Shokri. “On the Privacy Risks of Algorithmic Fairness”. In: *IEEE European Symposium on Security and Privacy*. 2021 (cit. on p. 27).

[CK21] Sapana Chaudhary and Dileep M. Kalathil. “Safe Online Convex Optimization with Unknown Linear Safety Constraints”. In: *AAAI Conference on Artificial Intelligence*. 2021 (cit. on p. 96).

[CDL20] Shuchi Chawla, Nikhil R. Devanur, and Thodoris Lykouris. “Static pricing for multi-unit prophet inequalities”. In: *Workshop on Internet and Network Economics*. 2020 (cit. on p. 195).

[Cha+10] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. “Multi-parameter mechanism design and sequential posted pricing”. In: *Behavioral and Quantitative Game Theory*. 2010 (cit. on p. 195).

[CH23] Chen and J. Hooker. “A guide to formulating fairness in an optimization model”. In: *Annals of Operations Research* (2023), pp. 1–39 (cit. on p. 16).

[CGS23] Tianrui Chen, Aditya Gangrade, and Venkatesh Saligrama. “Doubly-Optimistic Play for Safe Linear Bandits”. In: *ArXiv : 2209.13694* (2023) (cit. on pp. 96, 99, 103).

[CGS22] Tianrui Chen, Aditya Gangrade, and Venkatesh Saligrama. “Strategies for Safe Multi-Armed Bandits with Logarithmic Regret and Risk”. In: *International Conference on Machine Learning*. 2022 (cit. on p. 95).

[Che+19] Yifang Chen, Alex Cuellar, Haipeng Luo, et al. “Fair Contextual Multi-Armed Bandits: Theory and Experiments”. In: *Conference on Uncertainty in Artificial Intelligence*. 2019 (cit. on p. 95).

[Che+18] Yiling Chen, Nicole Immorlica, Brendan Lucier, Vasilis Syrgkanis, and Juba Ziani. “Optimal Data Acquisition for Statistical Estimation”. In: *Proceedings of the 2018 ACM Conference on Economics and Computation*. Ithaca, NY, USA, 2018 (cit. on p. 27).

[CZ19] Yiling Chen and Shuran Zheng. “Prior-Free Data Acquisition for Accurate Statistical Estimation”. In: *Proceedings of the ACM Conference on Economics and Computation*. 2019 (cit. on p. 27).

[Chi+19] Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. “Matroids, Matchings, and Fairness”. In: *International Conference on Artificial Intelligence and Statistics*. 2019 (cit. on pp. 257, 260, 273).

[Cho17] Alexandra Chouldechova. “Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments”. In: *Big Data* 5.2 (2017) (cit. on pp. 25, 313).

[Cho16] Alexandra Chouldechova. “Fair prediction with disparate impact: A study of bias in recidivism prediction instruments”. In: *Big data* 5 2 (2016), pp. 153–163 (cit. on p. 19).

[Cho+71] Y.S. Chow, Herbert Robbins, Siegmund David Herbert, and Siegmund David. “Great Expectations: The Theory of Optimal Stopping”. In: *Royal Statistical Society. Journal. Series A: General* (1971) (cit. on pp. 137, 192).

[Chu16] Youngsub Chun. *Fair queueing*. Springer, 2016 (cit. on p. 273).

[Chz+23] Evgenii Chzhen, Christophe Giraud, Z. Li, and Gilles Stoltz. “Small Total-Cost Constraints in Contextual Bandits with Knapsacks, with Application to Fairness”. In: *ArXiv : 2305.15807* (2023) (cit. on p. 95).

[Cla+20] Houston Claure, Yifang Chen, Jignesh Modi, Malte F. Jung, and Stefanos Nikolaidis. “Multi-Armed Bandits with Fairness Constraints for Distributing Resources to Human Teammates”. In: *IEEE International Conference on Human-Robot Interaction* (2020) (cit. on p. 95).

[Cle66] T. Anne Cleary. *Test Bias: Validity of the Scholastic Aptitude Test for Negro and White Students in Integrated Colleges*. Research Bulletin (RB-66-31). Found SAT slightly over-predictive for Black students in one of three colleges. Princeton, NJ: Educational Testing Service, June 1966 (cit. on p. 18).

[CS98] Melvyn G. Coles and Eric Smith. “Marketplaces and Matching”. In: *International Economic Review* 39 (1998), pp. 239–254 (cit. on p. 257).

[Col+20] Riccardo Colini-Baldeschi, Roberto Cominetti, Panayotis Mertikopoulos, and Marco Scarsini. “When is selfish routing bad? The price of anarchy in light and heavy traffic”. In: *Operations Research* 68.2 (2020), pp. 411–434 (cit. on p. 266).

[Com23] European Comission. “Commission proposes new measures on skills and talent to help address critical labour shortages”. In: (2023) (cit. on p. 257).

[Con+19] Vincent Conitzer, Rupert Freeman, Nisarg Shah, and Jennifer Wortman Vaughan. “Group Fairness for the Allocation of Indivisible Goods”. In: *AAAI Conference on Artificial Intelligence*. 2019 (cit. on p. 272).

[Coo+22] A. Feder Cooper, Samir Passi, Monica Lee, et al. “Accountability in an Algorithmic Society”. In: *FaccT* (2022) (cit. on p. 1).

[Cor+19a] Jose Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. “Recent developments in prophet inequalities”. In: *SIGecom Exch.* (2019) (cit. on pp. 137, 142, 160, 187).

[CC23] José Correa and Andrés Cristi. “A Constant Factor Prophet Inequality for Online Combinatorial Auctions”. In: *ACM Symposium on Theory of Computing* (2023) (cit. on p. 195).

[Cor+19b] José Correa, Paul Dütting, Felix A. Fischer, and Kevin Schewior. "Prophet Inequalities for I.I.D. Random Variables from an Unknown Distribution". In: *ACM Conference on Economics and Computation* (2019) (cit. on p. 196).

[Cor+17] José Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. "Posted Price Mechanisms for a Random Stream of Customers". In: *ACM Conference on Economics and Computation*. 2017 (cit. on pp. 138–140, 143–146, 148, 149, 160, 165, 171, 196).

[Cor+21] José R. Correa, Andrés Cristi, Paul Dütting, and Ashkan Norouzi-Fard. "Fairness and Bias in Online Selection". In: *International Conference on Machine Learning*. 2021 (cit. on p. 311).

[DG77] Claude D'Aspremont and Louis Gevers. "Equity and the Informational Basis of Collective Choice". In: *The Review of Economic Studies* 44.2 (1977), pp. 199–209 (cit. on pp. 16, 310).

[Dal20a] Hugh Dalton. "The Measurement of the Inequality of Incomes". In: *Economic Journal* 30.119 (1920), pp. 348–361 (cit. on p. 17).

[Dal20b] Hugh Dalton. "The Measurement of the Inequality of Incomes". In: *The Economic Journal* 30.119 (1920), pp. 348–361 (cit. on p. 203).

[Dar71] Richard B. Darlington. "Another Look at "Cultural Fairness"". In: *Journal of Educational Measurement* 8.2 (1971). Critically examines four definitions of cultural fairness and proposes a "cultural optimality" trade-off framework, pp. 71–82 (cit. on p. 18).

[Das18] Jeffrey Dastin. *Amazon scraps secret AI recruiting tool that showed bias against women*. 2018 (cit. on pp. 1, 25).

[Deb54] Gérard Debreu. "Valuation Equilibrium and Pareto Optimum". In: *Proceedings of the National Academy of Sciences* 40.7 (July 1954), pp. 588–592 (cit. on p. 1).

[DKT23] David Delacrétaz, Scott Duke Kominers, and Alexander Teytelboym. "Matching Mechanisms for Refugee Resettlement". In: *American Economic Review* (2023) (cit. on p. 257).

[DP94] Xiaotie Deng and Christos H. Papadimitriou. "On the Complexity of Cooperative Solution Concepts". In: *Mathematics of Operations Research* 19.2 (1994), pp. 257–266 (cit. on p. 14).

[Den+23] Yuan Deng, Negin Golrezaei, Patrick Jaillet, Jason Cheuk Nam Liang, and Vahab Mirrokni. "Multi-channel Autobidding with Budget and ROI Constraints". In: *International Conference on Machine Learning*. 2023 (cit. on p. 96).

[DG78] Robert Deschamps and Louis Gevers. "Leximin and utilitarian rules: A joint characterization". In: *Journal of Economic Theory* 17 (1978), pp. 143–163 (cit. on p. 310).

[Dev+23] Siddartha Devic, David Kempe, Vatsal Sharan, and Aleksandra Korolova. "Fairness in matching under uncertainty". In: *ICML*. 2023 (cit. on p. 273).

[DPS14] John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. “Price of fairness in kidney exchange”. In: *Adaptive Agents and Multi-Agent Systems*. 2014 (cit. on p. 272).

[Dic+18] John P. Dickerson, Karthik Abinav Sankararaman, Aravind Srinivasan, and Pan Xu. “Balancing Relevance and Diversity in Online Bipartite Matching via Submodularity”. In: *AAAI Conference on Artificial Intelligence*. 2018 (cit. on p. 25).

[Din+20] Dongsheng Ding, Xiaohan Wei, Zhuoran Yang, Zhaoran Wang, and Mihailo R. Jovanovi’c. “Provably Efficient Safe Exploration via Primal-Dual Policy Optimization”. In: *International Conference on Artificial Intelligence and Statistics*. 2020 (cit. on p. 96).

[DS61] Lester E. Dubins and Edwin H. Spanier. “How to Cut a Cake Fairly”. In: *American Mathematical Monthly* 68.1 (1961), pp. 1–17 (cit. on p. 10).

[DKA21] Piotr Dworczak, Scott Duke Kominers, and Mohammad Akbarpour. “Redistribution through markets”. In: *Econometrica* 89.4 (2021), pp. 1665–1698 (cit. on p. 204).

[Dwo+11] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard S. Zemel. “Fairness through awareness”. In: *ArXiv* abs/1104.3913 (2011) (cit. on p. 20).

[DR14] Cynthia Dwork and Aaron Roth. “The Algorithmic Foundations of Differential Privacy”. In: *Found. Trends Theor. Comput. Sci.* (2014) (cit. on pp. 40, 77).

[EOS07] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. “Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords”. In: *American Economic Review* 97.1 (2007), pp. 242–259 (cit. on p. 1).

[Edm70] Jack Edmonds. “Submodular functions, matroids, and certain polyhedra”. In: *Combinatorial Structures and Their Applications* (1970), pp. 69–87 (cit. on pp. 262, 276, 279).

[EMP20] Yonathan Efroni, Shie Mannor, and Matteo Pirotta. “Exploration-Exploitation in Constrained MDPs”. In: *ArXiv : 2003.02189* (2020) (cit. on p. 96).

[Ehs+17] Soheil Ehsani, Mohammad Taghi Hajiaghayi, Thomas Kesselheim, and Sahil Singla. “Prophet Secretary for Combinatorial Auctions and Matroids”. In: *ACM-SIAM Symposium on Discrete Algorithms*. 2017 (cit. on p. 195).

[El +20] Marwa El Halabi, Slobodan Mitrović, Ashkan Norouzi-Fard, Jakab Tardos, and Jakub M Tarnawski. “Fairness in streaming submodular maximization: Algorithms and hardness”. In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 13609–13622 (cit. on p. 273).

[El +24] Marwa El Halabi, Jakub Tarnawski, Ashkan Norouzi-Fard, and Thuy-Duong Vuong. “Fairness in Submodular Maximization over a Matroid Constraint”. In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 2024, pp. 1027–1035 (cit. on p. 273).

[Eme+20] Vitalii Emelianov, Nicolas Gast, Krishna P. Gummadi, and Patrick Loiseau. "On Fair Selection in the Presence of Implicit Variance". In: *Proceedings of the 21st ACM Conference on Economics and Computation*. Virtual Event, Hungary, 2020 (cit. on p. 25).

[Esf+19] Hossein Esfandiari, Mohammad Taghi Hajiaghayi, Brendan Lucier, and Michael Mitzenmacher. "Prophets, Secretaries, and Maximizing the Probability of Choosing the Best". In: *International Conference on Artificial Intelligence and Statistics*. 2019 (cit. on p. 196).

[Esf+17] Hossein Esfandiari, MohammadTaghi Hajiaghayi, Vahid Liaghat, and Morteza Monemizadeh. "Prophet Secretary". In: *SIAM Journal on Discrete Mathematics* (2017) (cit. on p. 195).

[Esm+22] Seyed-Alireza Esmaeili, Sharmila Duppala, Vedant Nanda, Aravind Srinivasan, and John P. Dickerson. "Rawlsian Fairness in Online Bipartite Matching: Two-sided, Group, and Individual". In: *Adaptive Agents and Multi-Agent Systems*. 2022 (cit. on p. 273).

[EF99] Péter Esö and Gábor Futó. "Auction design with a risk averse seller". In: *Economics Letters* 65.1 (1999), pp. 71–74 (cit. on p. 205).

[FK74] Allan Feldman and Alan Kirman. "Fairness and Envy". In: *The American Economic Review* 64.6 (1974), pp. 995–1005 (cit. on p. 210).

[Fer89] Thomas S. Ferguson. "Who Solved the Secretary Problem". In: *Statistical Science* (1989) (cit. on p. 137).

[Fio+22] Ferdinando Fioretto, Cuong Tran, Pascal Van Hentenryck, and Keyu Zhu. "Differential Privacy and Fairness in Decisions and Learning Tasks: A Survey". In: *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*. Survey Track. 2022 (cit. on p. 27).

[Fis65] Marek Fisz. *Probability Theory and Mathematical Statistics*. Wiley, New York, 3rd edition, 1965, pp. 372–377 (cit. on p. 251).

[FJ17] arthur Flajolet and Patrick Jaillet. "Real-Time Bidding with Side Information". In: *Advances in Neural Information Processing Systems*. 2017 (cit. on pp. 45, 46).

[FSV20] Rupert Freeman, Nisarg Shah, and Rohit Vaish. "Best of Both Worlds: Ex-Ante and Ex-Post Fairness in Resource Allocation". In: *Proceedings of the 21st ACM Conference on Economics and Computation* (2020) (cit. on p. 263).

[Fre+19] Rupert Freeman, Sujoy Sikdar, Rohit Vaish, and Lirong Xia. "Equitable Allocations of Indivisible Goods". In: *International Joint Conference on Artificial Intelligence*. 2019 (cit. on p. 9).

[Fre+23] Danielle N. Freund, Thodoris Lykouris, Elisabeth Paulson, Bradley Sturt, and Wen-Yu Weng. "Group fairness in dynamic refugee assignment". In: *ACM Conference on Economics and Computation*. 2023 (cit. on pp. 257, 272).

[FK16] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2016 (cit. on pp. 306, 307).

[GS62] David Gale and Lloyd S Shapley. "College admissions and the stability of marriage". In: *The American Mathematical Monthly* 69.1 (1962), pp. 9–15 (cit. on p. 1).

[GS13] David Gale and Lloyd S. Shapley. "College Admissions and the Stability of Marriage". In: *The American Mathematical Monthly* 120 (2013), pp. 386–391 (cit. on p. 257).

[GS58] George Gamow and Marvin Stern. *Puzzle Math*. 1958 (cit. on p. 10).

[Gel+20] Lodewijk L. Gelauff, Ashish Goel, Kamesh Munagala, and Sravya Yandamuri. "Advertising for Demographically Fair Outcomes". In: *ArXiv* (2020). Available as arXiv:2006.03983 (cit. on p. 25).

[GMS23] Giordano Giambartolomei, Frederik Mallmann-Trenn, and Raimundo Saona. "Prophet Inequalities: Separating Random Order from Order Selection". In: *ArXiv* (2023) (cit. on p. 195).

[Gil59] Donald B. Gillies. "3. Solutions to General Non-Zero-Sum Games". In: *Contributions to the Theory of Games, Volume IV*. Ed. by Albert William Tucker and Robert Duncan Luce. Princeton: Princeton University Press, 1959, pp. 47–86 (cit. on p. 14).

[Gin21] Corrado Gini. "Measurement of Inequality of Incomes". In: *The Economic Journal* 31.121 (1921), pp. 124–126 (cit. on p. 203).

[Gin12a] Corrado Gini. *Variabilità e Mutabilità*. Original introduction of the Gini coefficient. Bologna, Italy: C. Cuppini, 1912 (cit. on p. 17).

[Gin12b] Corrao Gini. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche. [Fasc. I.]* Studi economico-giuridici pubblicati per cura della facoltà di Giurisprudenza della R. Università di Cagliari. Tipogr. di P. Cuppini, 1912 (cit. on p. 203).

[GO03] Jacob K. Goeree and Theo Offerman. "Competitive bidding in auctions with private and common values". In: *The Economic Journal* 113.489 (2003), pp. 598–613 (cit. on pp. 202, 248).

[GLR20] Thibaut Le Gouic, Jean-Michel Loubes, and Philippe Rigollet. "Projection to Fairness in Statistical Learning." In: *arXiv* (2020) (cit. on p. 273).

[GSM21] GSMA. *Auction Best Practice GSMA Public Policy Position*. Tech. rep. GSMA Association, 2021 (cit. on p. 202).

[Gua23] Guardian. "Facebook and Instagram to restrict advertisers' access to teenagers' data". In: *The Guardian* (2023) (cit. on pp. 4, 25).

[Haa19] Christian Haas. "The Price of Fairness - A Framework to Explore Trade-Offs in Algorithmic Fairness". In: *International Conference on Interaction Sciences*. 2019 (cit. on p. 273).

[HKS07] Mohammad Taghi Hajiaghayi, Robert D. Kleinberg, and Tuomas Sandholm. "Automated Online Mechanism Design and Prophet Inequalities". In: *AAAI Conference on Artificial Intelligence*. 2007 (cit. on p. 195).

[Ham76] Peter J. Hammond. “Equity, Arrow’s Conditions, and Rawls’ Difference Principle”. In: *Econometrica* 44.4 (1976), pp. 793–804 (cit. on p. 16).

[Har25] Elfarouk Harb. “New Prophet Inequalities via Poissonization and Sharding”. In: *ACM-SIAM Symposium on Discrete Algorithms*. 2025 (cit. on p. 196).

[Har+16a] Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou, and Mary Wootters. “Strategic Classification”. In: *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science (ITCS)*. ACM, 2016, pp. 111–122 (cit. on p. 312).

[Har+16b] Moritz Hardt, Eric Price, Eric Price, and Nati Srebro. “Equality of Opportunity in Supervised Learning”. In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett. Vol. 29. Curran Associates, Inc., 2016 (cit. on p. 19).

[Har+16c] Moritz Hardt, Eric Price, Eric Price, and Nati Srebro. “Equality of Opportunity in Supervised Learning”. In: *Advances in Neural Information Processing Systems*. Vol. 29. 2016 (cit. on pp. 258, 272).

[HPS16] Moritz Hardt, Eric Price, and Nathan Srebro. “Equality of Opportunity in Supervised Learning”. In: *Advances in Neural Information Processing Systems*. 2016 (cit. on pp. 25, 29, 42).

[HM05] John William Hatfield and Paul R Milgrom. “Matching with contracts”. In: *American Economic Review* 95.4 (2005), pp. 913–935 (cit. on p. 1).

[Haz22] Elad Hazan. *Introduction to Online Convex Optimization*. 2nd. The MIT Press, 2022 (cit. on pp. 36, 59).

[HK18] Hoda Heidari and Andreas Krause. “Preventing Disparate Treatment in Sequential Decision Making”. In: *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*. 2018 (cit. on p. 25).

[HH02] Jürgen Herzog and Takayuki Hibi. “Discrete polymatroids”. In: *Journal of Algebraic Combinatorics* 16 (2002), pp. 239–268 (cit. on pp. 264, 309).

[HH11] Jürgen Herzog and Takayuki Hibi. *Monomial Ideals*. Springer, 2011 (cit. on p. 279).

[Hil83] T. P. Hill. “Prophet Inequalities and Order Selection in Optimal Stopping Problems”. In: *Proceedings of the American Mathematical Society* (1983) (cit. on p. 137).

[HK82] T. P. Hill and Robert P. Kertz. “Comparisons of Stop Rule and Supremum Expectations of I.I.D. Random Variables”. In: *The Annals of Probability* (1982) (cit. on pp. 137, 138, 159, 190, 192).

[HK83] T. P. Hill and Robert P. Kertz. “Stop Rule Inequalities for Uniformly Bounded Sequences of Random Variables”. In: *Transactions of the American Mathematical Society* (1983) (cit. on p. 137).

[HT23] Pär Holmberg and Thomas Tangerås. “A Survey of Capacity Mechanisms: Lessons for the Swedish Electricity Market”. In: *The Energy Journal* 44.6 (2023), pp. 275–304 (cit. on p. 202).

[HT11] Junya Honda and Akimichi Takemura. “An asymptotically optimal policy for finite support models in the multiarmed bandit problem”. In: *Machine Learning* (2011) (cit. on p. 98).

[Hos+23] Hadi Hosseini, Zhiyi Huang, Ayumi Igarashi, and Nisarg Shah. “Class fairness in online matching”. In: *AAAI*. 2023 (cit. on p. 273).

[HIL19] L. Elisa Celis Hu, Nicole Immorlica, and Brendan Lucier. “The Disparate Effects of Strategic Manipulation”. In: *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*. ACM, 2019, pp. 499–500 (cit. on p. 312).

[HTA23] Spencer Hutchinson, Berkay Turan, and Mahnoosh Alizadeh. “Exploiting Problem Geometry in Safe Linear Bandits”. In: *ArXiv : 2308.15006* (2023) (cit. on p. 95).

[HZ79] Aanund Hylland and Richard J. Zeckhauser. “The Efficient Allocation of Individuals to Positions”. In: *Journal of Political Economy* 87 (1979), pp. 293–314 (cit. on p. 263).

[Imm+22] Nicole Immorlica, Karthik Sankararaman, Robert Schapire, and Aleksandrs Slivkins. “Adversarial Bandits with Knapsacks”. In: *Journal of the ACM* (2022) (cit. on p. 95).

[IJS11] Krishnamurthy Iyer, Ramesh Johari, and Mukund Sundararajan. “Mean field equilibria of dynamic auctions with learning: A dynamic revenue equivalence theorem”. In: *49th Annual Allerton Conference on Communication, Control, and Computing* (2011) (cit. on p. 45).

[Jab+17] Shahin Jabbari, Matthew Joseph, Michael Kearns, Jamie Morgenstern, and Aaron Roth. “Fairness in Reinforcement Learning”. In: *Proceedings of the 34th International Conference on Machine Learning*. 2017 (cit. on p. 25).

[Jag+19] Matthew Jagielski, Michael Kearns, Jieming Mao, et al. “Differentially Private Fair Learning”. In: *Proceedings of the 36th International Conference on Machine Learning*. 2019 (cit. on p. 27).

[JCH98] Raj Jain, D. M. Chiu, and William R. Hawe. “A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems”. In: *ArXiv* (1998) (cit. on p. 17).

[Jen+16] Rodolphe Jenatton, Jim Huang, Dominik Csiba, and Cedric Archambeau. *Online optimization and regret guarantees for non-additive long-term constraints*. Available as arXiv:1602.05394. 2016 (cit. on pp. 27, 55).

[JHA15] Rodolphe Jenatton, Jim C. Huang, and C. Archambeau. “Adaptive Algorithms for Online Convex Optimization with Long-term Constraints”. In: *International Conference on Machine Learning*. 2015 (cit. on p. 96).

[JMZ21] Jiashuo Jiang, Will Ma, and Jiawei Zhang. “Tight Guarantees for Multi-unit Prophet Inequalities and Online Stochastic Knapsack”. In: *ACM-SIAM Symposium on Discrete Algorithms*. 2021 (cit. on p. 195).

[JMZ22] Jiashuo Jiang, Will Ma, and Jiawei Zhang. “Tightness without Counterexamples: A New Approach and New Results for Prophet Inequalities”. In: *ACM Conference on Economics and Computation* (2022) (cit. on pp. 138, 141, 151, 159, 195).

[Jos+16] Matthew Joseph, Michael Kearns, Jamie Morgenstern, and Aaron Roth. “Fairness in Learning: Classic and Contextual Bandits”. In: *Advances in Neural Information Processing Systems*. 2016 (cit. on p. 25).

[Kah+01] Alfred E. Kahn, Peter C. Cramton, Robert H. Porter, and Richard D. Tabors. “Uniform Pricing or Pay-as-Bid Pricing: A Dilemma for California and Beyond”. In: *The Electricity Journal* 14.6 (2001), pp. 70–79 (cit. on p. 201).

[KS75] Ehud Kalai and Meir Smorodinsky. “Other solutions to Nash’s bargaining problem”. In: *Econometrica* 43.3 (1975), pp. 513–518 (cit. on pp. 13, 258, 272).

[KK23] Yuichiro Kamada and Fuhito Kojima. “Fair Matching under Constraints: Theory and Applications”. In: *Review of Economic Studies* (2023) (cit. on p. 273).

[KAS11] Toshihiro Kamishima, Shotaro Akaho, and Jun Sakuma. “Fairness-aware Learning through Regularization Approach”. In: *IEEE 11th International Conference on Data Mining Workshops*. 2011 (cit. on p. 29).

[KRZ18] Sampath Kannan, Aaron Roth, and Juba Ziani. “Downstream Effects of Affirmative Action”. In: *Proceedings of the Conference on Fairness, Accountability, and Transparency* (2018) (cit. on p. 15).

[Kas23] Bernhard Kasberger. “When can auctions maximize post-auction welfare?” In: *International Journal of Industrial Organization* 89 (2023), p. 102972 (cit. on p. 202).

[Kea+17] Michael Kearns, Seth Neel, Aaron Roth, and Zhiwei Steven Wu. “Preventing Fairness Gerrymandering: Auditing and Learning for Subgroup Fairness”. In: *ArXiv* abs/1711.05144 (2017) (cit. on p. 19).

[Kea+18] Michael Kearns, Seth Neel, Aaron Roth, and Zhiwei Steven Wu. “Preventing Fairness Gerrymandering: Auditing and Learning for Subgroup Fairness”. In: *Proceedings of the 35th International Conference on Machine Learning*. Vol. 80. PMLR, 2018, pp. 2564–2572 (cit. on p. 265).

[Ken85] D. P. Kennedy. “Optimal Stopping of Independent Random Variables and Maximizing Prophets”. In: *Annals of Probability* (1985) (cit. on pp. 138, 140, 149, 195).

[Ken87] D.P. Kennedy. “Prophet-type inequalities for multi-choice optimal stopping”. In: *Stochastic Processes and their Applications* (1987) (cit. on pp. 141, 154, 195).

[Ker86a] Robert P Kertz. "Stop rule and supremum expectations of i.i.d. random variables: a complete comparison by conjugate duality". In: *J. Multivar. Anal.* (1986) (cit. on pp. 137, 138, 140, 151).

[Ker86b] Robert P. Kertz. "Comparison of Optimal Value and Constrained Maxima Expectations for Independent Random Variables". In: *Advances in Applied Probability* (1986) (cit. on pp. 138, 149).

[KMR18] Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. "Algorithmic fairness". In: *AEA Papers and Proceedings* 108 (2018), pp. 22–27 (cit. on p. 311).

[KR18] Jon Kleinberg and Manish Raghavan. "Selection Problems in the Presence of Implicit Bias". In: *9th Innovations in Theoretical Computer Science Conference*. 2018 (cit. on p. 25).

[KW12a] Robert Kleinberg and Seth Matthew Weinberg. "Matroid prophet inequalities". In: *ACM Symposium on Theory of Computing*. 2012 (cit. on p. 195).

[KW12b] Robert D. Kleinberg and S. Matthew Weinberg. "Matroid Prophet Inequalities". In: *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC)*. ACM, 2012, pp. 123–136 (cit. on p. 312).

[Kle98] Paul Klemperer. "Auctions with almost common values: The 'Wallet Game' and its applications". In: *European Economic Review* 42.3 (1998), pp. 757–769 (cit. on p. 202).

[Koj12] Fuhito Kojima. "School choice: Impossibilities for affirmative action". In: *Games and Economic Behavior* 75.2 (2012), pp. 685–693 (cit. on p. 15).

[KS77] Ulrich Krengel and Louis Sucheston. "Semiamarts and finite values". In: *Bulletin of the American Mathematical Society* (1977) (cit. on p. 137).

[Kri09] Vijay Krishna. *Auction Theory*. Academic Press, 2009 (cit. on pp. 206, 208, 209, 211, 232, 233).

[KD69] Robert L. Kruse and John J. Deely. "Joint continuity of monotonic functions". In: *The American Mathematical Monthly* 76.1 (1969), pp. 74–76 (cit. on p. 240).

[KK22] Raunak Kumar and Robert Kleinberg. "Non-monotonic Resource Utilization in the Bandits with Knapsacks Problem". In: *Advances in Neural Information Processing Systems*. 2022 (cit. on p. 95).

[LT19a] Anja Lambrecht and Catherine E. Tucker. "Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads". In: *Manag. Sci.* 65 (2019), pp. 2966–2981 (cit. on p. 257).

[LT19b] Anja Lambrecht and E. Tucker Catherine. "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads". In: *Management Science*. 2019 (cit. on p. 25).

[Lam+19] Alex Lamy, Ziyuan Zhong, Aditya K Menon, and Nakul Verma. "Noise-tolerant fair classification". In: *Advances in Neural Information Processing Systems*. 2019 (cit. on p. 27).

[Lar+16] Jeff Larson, Surya Mattu, Lauren Kirchner, and Julia Angwin. *How We Analyzed the COMPAS Recidivism Algorithm*. 2016 (cit. on pp. 1, 25).

[LS20a] Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020 (cit. on pp. 32, 35, 37–39, 59, 60, 75–77, 86, 89).

[LS20b] Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020 (cit. on pp. 92, 123).

[LSY22] Xiaocheng Li, Chunlin Sun, and Yinyu Ye. "Simple and Fast Algorithm for Binary Integer and Online Linear Programming". In: *Mathematical Programming* (2022) (cit. on p. 27).

[LSY21] Xiaocheng Li, Chunlin Sun, and Yinyu Ye. "The Symmetry between Arms and Knapsacks: A Primal-Dual Approach for Bandits with Knapsacks". In: *International Conference on Machine Learning*. Proceedings of Machine Learning Research. 2021 (cit. on pp. 95, 96).

[LY21] Xiaocheng Li and Yinyu Ye. "Online Linear Programming: Dual Convergence, New Algorithms, and Regret Bounds". In: *Operations Research* (2021) (cit. on p. 27).

[LLZ23] Jason Cheuk Nam Liang, Haihao Lu, and Baoyu Zhou. "Online Ad Procurement in Non-stationary Autobidding Worlds". In: *ArXiv : 2307.05698* (2023) (cit. on p. 96).

[Lip+04] Richard J Lipton, Evangelos Markakis, Elchanan Mossel, and Amin Saberi. "On approximately fair allocations of indivisible goods". In: *ACM Conference on Electronic Commerce*. 2004 (cit. on p. 272).

[LMC18] Zachary Lipton, Julian McAuley, and Alexandra Chouldechova. "Does mitigating ML's impact disparity require treatment disparity?" In: *Advances in Neural Information Processing Systems*. 2018 (cit. on p. 25).

[LKB+17] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, et al. "A survey on deep learning in medical image analysis". In: *Medical Image Analysis* 42 (2017), pp. 60–88 (cit. on p. 1).

[Liu+21a] Allen Liu, Renato Paes Leme, Martin Pál, Jon Schneider, and Balasubramanian Sivan. "Variable Decomposition for Prophet Inequalities and Optimal Ordering". In: *ACM Conference on Economics and Computation*. 2021 (cit. on pp. 139, 149, 153, 169, 174).

[Liu+18] Lydia T. Liu, Sarah Dean, Erik Rolf, Max Simchowitz, and Moritz Hardt. "Delayed Impact of Fair Machine Learning". In: *Proceedings of the 35th International Conference on Machine Learning (ICML)*. Vol. 80. Proceedings of Machine Learning Research. PMLR, 2018, pp. 3150–3158 (cit. on p. 312).

[Liu+22] Qingsong Liu, Weihang Xu, Si-Yi Wang, and Zhixuan Fang. "Combinatorial Bandits with Linear Constraints: Beyond Knapsacks and Fairness". In: *Neural Information Processing Systems*. 2022 (cit. on p. 95).

[Liu+21b] Xin Liu, Bin Li, P. Shi, and Lei Ying. "An Efficient Pessimistic-Optimistic Algorithm for Stochastic Linear Bandits with General Constraints". In: *Neural Information Processing Systems*. 2021 (cit. on p. 95).

[Liu+21c] Xin Liu, Bin Li, Pengyi Shi, and Lei Ying. "An Efficient Pessimistic-Optimistic Algorithm for Stochastic Linear Bandits with General Constraints". In: *Advances in Neural Information Processing Systems*. 2021 (cit. on p. 27).

[Lor05] Max O. Lorenz. "Methods of Measuring the Concentration of Wealth". In: *Publications of the American Statistical Association* 9.70 (1905), pp. 209–219 (cit. on p. 203).

[LC23] Orestis Loukas and Ho-Ryun Chung. "Demographic parity: Mitigating biases in real-world data". In: *arXiv preprint arXiv:2309.17347* (2023) (cit. on p. 272).

[MX20] Will Ma and Pan Xu. "Group-level Fairness Maximization in Online Bipartite Matching". In: *Adaptive Agents and Multi-Agent Systems*. 2020 (cit. on p. 311).

[MXX21] Will Ma, Pan Xu, and Yifan Xu. "Fairness Maximization among Offline Agents in Online-Matching Markets". In: 2021 (cit. on p. 273).

[MTY09] Shie Mannor, John N. Tsitsiklis, and Jia Yuan Yu. "Online Learning with Sample Path Constraints". In: *Journal of Machine Learning Research* (2009) (cit. on p. 96).

[MTL20] Daniel Marszalec, Alexander Teytelboym, and Sanna Laksá. *EPIC Fail: How Below-Bid Pricing Backfires in Multiunit Auctions*. Working Paper. 2020 (cit. on p. 207).

[Meh13] Aranyak Mehta. "Online Matching and Ad Allocation". In: *Found. Trends Theor. Comput. Sci.* (2013) (cit. on pp. 137, 257).

[MW21] Aditya Krishna Menon and Robert C. Williamson. "The cost of fairness in binary classification". In: *Conference on Fairness, Accountability, and Transparency*. 2021 (cit. on p. 273).

[MS02] Paul Milgrom and Ilya Segal. "Envelope Theorems for Arbitrary Choice Sets". In: *Econometrica* 70.2 (2002), pp. 583–601 (cit. on pp. 202, 208).

[MW98] Jeonghoon Mo and Jean C. Walrand. "Fair end-to-end window-based congestion control". In: *TNET*. 1998 (cit. on p. 16).

[ML22a] Mathieu Molina and Patrick Loiseau. "Bounding and Approximating Intersectional Fairness through Marginal Fairness". In: *Advances in Neural Information Processing Systems*. 2022 (cit. on p. 25).

[ML22b] Mathieu Molina and Patrick Loiseau. "Bounding and Approximating Intersectional Fairness through Marginal Fairness". In: *Advances in Neural Information Processing Systems*. Vol. 35. 2022, pp. 16796–16807 (cit. on p. 265).

[MS22] Faidra Monachou and Ana-Andreea Stoica. “Fairness and equity in resource allocation and decision-making: an annotated reading list”. In: *ACM SIGecom Exchanges* 20.1 (2022), pp. 64–66 (cit. on p. 2).

[Mor+21] Ahmadreza Moradipari, Sanae Amani, Mahnoosh Alizadeh, and Christos Thrampoulidis. “SAFE LINEAR BANDITS”. In: *Annual Conference on Information Sciences and Systems*. 2021 (cit. on p. 95).

[MPT94] Rajeev Motwani, Steven Phillips, and Eric Torng. “Nonclairvoyant scheduling”. In: *Theoretical Computer Science* (1994) (cit. on p. 137).

[Mou03] Hervé Moulin. “Fair division and collective welfare”. In: 2003 (cit. on p. 9).

[Mou04] Hervé Moulin. *Fair division and collective welfare*. MIT press, 2004 (cit. on pp. 203, 210).

[Mye91] Roger B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991 (cit. on p. 12).

[Mye81] Roger B. Myerson. “Optimal Auction Design”. In: *Mathematics of Operations Research* (1981) (cit. on p. 226).

[Nag85] John Nagle. *On Packet Switches With Infinite Storage*. RFC 970, Request for Comments. Dec. 1985 (cit. on p. 16).

[Nas50] John F. Jr. Nash. “The Bargaining Problem”. In: *Econometrica* 18.2 (1950), pp. 155–162 (cit. on p. 12).

[NT20] Milad Nasr and Michael Carl Tschantz. “Bidding Strategies with Gender Nondiscrimination Constraints for Online Ad Auctions”. In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 2020 (cit. on p. 25).

[New14] New Zealand Electricity Authority. *Transmission pricing methodology review: Beneficiaries-pay options*. Tech. rep. New Zealand Electricity Authority, Wellington, 2014 (cit. on p. 202).

[NPP17] Gaia Nicosia, Andrea Pacifici, and Ulrich Pferschy. “Price of Fairness for allocating a bounded resource”. In: *European Journal of Operational Research* (2017), pp. 933–943 (cit. on p. 272).

[Nou95] Charles N. Noussair. “Equilibria in a multi-object uniform price sealed bid auction with multi-unit demands”. In: *Economic Theory* 5 (1995), pp. 337–351 (cit. on p. 225).

[OCo25] Sarah O’Connor. “The AI arms race in hiring is a huge mess for everyone”. In: *Financial Times* (May 6, 2025). Online; accessed June 2025 (cit. on p. 1).

[OEC21] OECD. *OECD Sovereign Borrowing Outlook 2021*. 2021, p. 90 (cit. on p. 201).

[Ora19] Francesco Orabona. *A Modern Introduction to Online Learning*. Available as arXiv:1912.13213. 2019 (cit. on pp. 59, 62).

[OR95] Martin J. Osborne and Ariel Rubinstein. “A Course in Game Theory”. In: MIT Press, 1995 (cit. on p. 308).

[Pac+21] Aldo Pacchiano, Mohammad Ghavamzadeh, Peter L. Bartlett, and Heinrich Jiang. "Stochastic Bandits with Linear Constraints". In: *International Conference on Artificial Intelligence and Statistics*. 2021 (cit. on pp. 95, 96).

[PV12] Mallesh M. Pai and Rakesh Vohra. *Auction Design with Fairness Concerns: Subsidies vs. Set-Asides*. Discussion Paper #1548. 2012 (cit. on pp. 201, 204).

[Pas15] Frank Pasquale. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press, 2015 (cit. on p. 1).

[Pat+21] Vishakha Patil, Ganesh Ghalme, Vineet Jagadeesan Nair, and Y. Narahari. "Achieving Fairness in the Stochastic Multi-armed Bandit Problem". In: *Journal of Machine Learning Research* (2021) (cit. on p. 95).

[PP19] John W. Patty and Elizabeth Maggie Penn. "Measuring Fairness, Inequality, and Big Data: Social Choice Since Arrow". In: *Annual Review of Political Science* 22.1 (2019), pp. 435–460 (cit. on p. 210).

[PR13] Vianney Perchet and Philippe Rigollet. "The Multi-Armed Bandit Problem with Covariates". In: *The Annals of Statistics* 41.2 (2013) (cit. on pp. 30, 38).

[Per+16] Vianney Perchet, Philippe Rigollet, Sylvain Chassang, and Erik Snowberg. "Batched bandit problems". In: (2016) (cit. on p. 98).

[PST22] Sebastian Perez-Salazar, Mohit Singh, and Alejandro Toriello. "The IID Prophet Inequality with Limited Flexibility". In: *ArXiv* (2022) (cit. on p. 138).

[Pig12a] Arthur C. Pigou. *Wealth and Welfare*. Original formulation of income-transfer principle. London: Macmillan, 1912 (cit. on p. 17).

[Pig12b] Arthur Cecil Pigou. *Wealth and welfare*. London: Macmillan and Company, limited, 1912 (cit. on p. 203).

[RAW71] JOHN RAWLS. *A Theory of Justice: Original Edition*. Harvard University Press, 1971 (cit. on p. 15).

[Ref23] United Nations High Commissioner for Refugees. *Resettlement Fact Sheet 2023*. 2023 (cit. on p. 257).

[RS81] John G. Riley and William F. Samuelson. "Optimal Auctions". In: *The American Economic Review* (1981) (cit. on p. 226).

[Roc70] Ralph Tyrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970 (cit. on p. 53).

[RSÜ04] Alvin E Roth, Tayfun Sönmez, and M Utku Ünver. "Kidney exchange". In: *Quarterly Journal of Economics* 119.2 (2004), pp. 457–488 (cit. on p. 1).

[Rou20] Tim Roughgarden. *Beyond the Worst-Case Analysis of Algorithms*. Cambridge University Press, 2020 (cit. on pp. 258, 273).

[Roz+22] Benedek Rozemberczki, Lauren Watson, Péter Bayer, et al. "The Shapley Value in Machine Learning". In: *ArXiv* abs/2202.05594 (2022) (cit. on p. 14).

[RWW20] Aviad Rubinstein, Jack Z. Wang, and S. Matthew Weinberg. "Optimal Single-Choice Prophet Inequalities from Samples". In: *Innovations in Theoretical Computer Science Conference*. 2020 (cit. on pp. 140, 196).

[RPD17] Keith Ruddell, Andy Philpott, and Anthony Downward. "Supply Function Equilibrium with Taxed Benefits". In: *Operations Research* 65.1 (2017), pp. 1–18 (cit. on pp. 202, 207).

[S S+21] Govind S. Sankar, Anand Louis, Meghana Nasre, and Prajakta Nimbhorkar. "Matchings with Group Fairness Constraints: Online and Offline Algorithms". In: *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*. 2021, pp. 377–383 (cit. on pp. 272, 273).

[Sam84] Ester Samuel-Cahn. "Comparison of Threshold Stop Rules and Maximum for Independent Nonnegative Random Variables". In: *Annals of Probability* (1984) (cit. on p. 137).

[SS20] Karthik Abinav Sankararaman and Aleksandrs Slivkins. "Bandits with Knapsacks beyond the Worst Case". In: *Neural Information Processing Systems*. 2020 (cit. on p. 95).

[Sch+03] Alexander Schrijver et al. *Combinatorial optimization: polyhedra and efficiency*. Vol. 24. 2. Springer, 2003 (cit. on pp. 257, 260, 273, 275).

[Sen17] Amartya Sen. *An Expanded Edition*. Cambridge, MA and London, England: Harvard University Press, 2017 (cit. on pp. 16, 272).

[SF73] Amartya Sen and James Foster. *On Economic Inequality*. Oxford University Press, 1973 (cit. on p. 203).

[Sha53] Lloyd S. Shapley. "A value for n-person games". In: *Contributions to the Theory of Games II*. Ed. by H. W. Kuhn and A. W. Tucker. Vol. 28. Annals of Mathematics Studies. Princeton University Press, 1953, pp. 307–317 (cit. on p. 14).

[Sha71] Lloyd S. Shapley. "Cores of Convex Games". In: *International Journal of Game Theory* 1 (1971), pp. 11–26 (cit. on p. 14).

[Sin23] Abhishek Sinha. "BanditQ: Fair Multi-Armed Bandits with Guaranteed Rewards per Arm". In: *Arxiv:2304.05219*. 2023 (cit. on pp. 93, 95, 106, 107, 132, 133, 339).

[Siv+21] Balasubramanian Sivan, Hedyeh Beyhaghi, Martin Pál, Negin Golrezaei, and Renato Paes Leme. "Improved Approximations for Posted Price and Second-price Mechanisms". In: *Operations Research* (2021) (cit. on p. 195).

[Sli19] Aleksandrs Slivkins. "Introduction to Multi-Armed Bandits". In: *Foundations and Trends® in Machine Learning* 12.1-2 (2019) (cit. on p. 38).

[SSF22] Aleksandrs Slivkins, Karthik Abinav Sankararaman, and Dylan J. Foster. "Contextual Bandits with Packing and Covering Constraints: A Modular Lagrangian Approach via Regression". In: *Annual Conference Computational Learning Theory*. 2022 (cit. on pp. 93, 95, 106, 109, 130, 131, 133, 339).

[Spe+18a] Till Speicher, Muhammad Ali, Giridhari Venkatadri, et al. "On the Potential for Discrimination in Online Targeted Advertising". In: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. 2018 (cit. on p. 25).

[Spe+18b] Till Speicher, Muhammad Ali, Giridhari Venkatadri, et al. "Potential for Discrimination in Online Targeted Advertising". In: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. Vol. 81. PMLR, 2018, pp. 5–19 (cit. on p. 257).

[Ste49] Hugo Steinhaus. "Sur la division pragmatique". In: *Econometrica: Journal of the Econometric Society* (1949), pp. 315–319 (cit. on p. 272).

[Ste48] Hugo Steinhaus. "The problem of fair division". In: *Econometrica* 16.1 (1948), pp. 101–104 (cit. on p. 9).

[TY23] Shaojie Tang and Jing Yuan. "Beyond submodularity: a unified framework of randomized set selection with group fairness constraints". In: *Journal of Combinatorial Optimization* 45 (2023), pp. 1–22 (cit. on p. 273).

[Tho33] William R Thompson. "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples". In: *Biometrika* (1933) (cit. on p. 98).

[Tho11] William Thomson. "Chapter Twenty-One - Fair Allocation Rules". In: ed. by Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura. Vol. 2. Handbook of Social Choice and Welfare. Elsevier, 2011, pp. 393–506 (cit. on p. 205).

[Tyl90] Tom R. Tyler. *Why People Obey the Law*. Princeton, NJ: Princeton University Press, 1990 (cit. on p. 3).

[US 24] U.S. Small Business Administration. *FY 2024 Goaling Guidelines*. Tech. rep. Office of Policy, Planning & Liaison, Office of Government Contracting & Business Development, 2024 (cit. on p. 201).

[Var74] Hal R Varian. "Equity, envy, and efficiency". In: *Journal of Economic Theory* 9.1 (1974), pp. 63–91 (cit. on p. 272).

[Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018 (cit. on pp. 78, 79).

[Vic61] William Vickrey. "Counterspeculation, Auctions, and Competitive Sealed Tenders". In: *The Journal of Finance* 16.1 (1961), pp. 8–37 (cit. on p. 211).

[VW02] Sivakumar Viswanathan and James J. D. Wang. "Market architecture: limit-order books versus dealership markets". In: *Journal of Financial Markets* 5.2 (2002), pp. 127–167 (cit. on p. 207).

[WZ02] James J. D. Wang and Jaime F. Zender. "Auctioning Divisible Goods". In: *Economic Theory* 19.4 (2002), pp. 673–705 (cit. on p. 207).

[Wan+21] Lequn Wang, Yiwei Bai, Wen Sun, and Thorsten Joachims. "Fairness of Exposure in Stochastic Bandits". In: *International Conference on Machine Learning*. 2021 (cit. on p. 95).

[Wel85] Dietrich Weller. “Fair division of a measurable space”. In: *Journal of Mathematical Economics* 14 (1985), pp. 5–17 (cit. on p. 272).

[Woo21] Kyle Woodward. *Mixed-Price Auctions for Divisible Goods*. Working paper. 2021 (cit. on p. 207).

[XLL20] Tengyu Xu, Yingbin Liang, and Guanghui Lan. “CRPO: A New Approach for Safe Reinforcement Learning with Convergence Guarantee”. In: *International Conference on Machine Learning*. 2020 (cit. on p. 96).

[You94] H. Peyton Young. *Equity: In Theory and Practice*. Princeton University Press, 1994 (cit. on p. 12).

[YNW17] Hao Yu, Michael J. Neely, and Xiaohan Wei. “Online Convex Optimization with Stochastic Constraints”. In: *Neural Information Processing Systems*. 2017 (cit. on p. 96).

[YT23] Jing Yuan and Shaojie Tang. “Group fairness in non-monotone submodular maximization”. In: *Journal of Combinatorial Optimization* 45 (2023), pp. 1–15 (cit. on p. 273).

[Zha+22] Tianxiang Zhao, Enyan Dai, Kai Shu, and Suhang Wang. “Towards Fair Classifiers Without Sensitive Attributes: Exploring Biases in Related Features”. In: *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*. Virtual Event, AZ, USA, 2022 (cit. on p. 27).

[ZB16] Li Zhou and Emma Brunskill. “Latent Contextual Bandits and Their Application to Personalized Recommendations for New Users”. In: *International Joint Conference on Artificial Intelligence*. 2016 (cit. on p. 93).

[ZJ22] Xingyu Zhou and Bo Ji. “On Kernelized Multi-Armed Bandits with Constraints”. In: *Advances in Neural Information Processing Systems*. 2022 (cit. on p. 95).

[Zon+16] Shi Zong, Hao Ni, Kenny Sung, et al. “Cascading Bandits for Large-Scale Recommendation Problems”. In: *Conference on Uncertainty in Artificial Intelligence*. 2016 (cit. on p. 93).

List of Figures

ECOLE
DOCTORALE
DE MATHEMATIQUES
HADAMARD

Titre : L'impact de l'équité dans les problèmes d'allocation de ressources

Mots clés : Equité, Théorie des Jeux, Algorithmes Séquentiels, Statistiques

Résumé : Les problèmes d'allocation de ressources, allant de l'affectation scolaire à la publicité en ligne en passant par les soins de santé, privilégient traditionnellement l'efficacité et la maximisation d'utilité, en négligeant souvent les considérations éthiques. L'introduction de contraintes d'équité dans des problèmes d'allocation classiques requiert de nouvelles méthodes algorithmiques, impacte les performances, et engendre de nouveaux compromis entre optimalité et respect des contraintes. Cette thèse étudie quatre modèles classiques d'allocation – apprentissage en ligne, inégalités de prophète, enchères, et allocation sous contraintes de matroïdes – sous l'angle de l'équité, et examine comment l'imposition de critères d'équités affecte la qualité de l'allocation à travers des bornes de regret, des ratios de compétition, des mesures d'inégalité et des notions de prix d'équité. Nous montrons que, dans certains cas, les compromis sont inévitables ; dans d'autres, la perte peut devenir arbitrairement faible asymptotiquement ; et dans des cas spécifiques, l'équité peut être atteinte sans aucun coût. En caractérisant quand et comment les contraintes d'équités influencent les résultats d'allocation, cette thèse apporte des éléments de réponse utiles tant aux concepteurs d'algorithmes qu'aux décideurs publics souhaitant mettre en œuvre des algorithmes équitables dans des contextes d'allocation.

Title : The impact of fairness in resource allocation

Keywords : Fairness, Online Algorithms, Game Theory, Statistics

Abstract : Resource allocation problems, ranging from school choice to online advertising and healthcare, traditionally prioritize efficiency and social welfare while often overlooking ethical concerns. Introducing fairness constraints into well-established allocation problems requires new algorithmic design principles, alters performance, and gives rise to new trade-offs between optimality and constraints satisfaction. This thesis studies four standard allocation models – online learning, prophet inequalities, auctions, and matroid-based allocation – with equity considerations, and explores how enforcing fairness impacts the allocation quality through regret bounds, competitive ratio, inequality measures, and price of fairness. We show that, in some cases, trade-offs are unavoidable; in others, the loss becomes arbitrarily small asymptotically; and in specific instances, fairness can be achieved at no cost. By characterizing when and how fairness constraints affect allocation outcomes, this thesis provides guidance to both algorithm designers and policy makers seeking to deploy fair algorithms in some allocation problem scenarios.

Institut Polytechnique de Paris
91120 Palaiseau, France