

Fair Multi-Agent Persuasion with Submodular Constraints

Yannan Bai ✉

Duke University, Durham, NC, USA

Kamesh Munagala ✉ 

Duke University, Durham, NC, USA

Yiheng Shen ✉ 

Duke University, Durham, NC, USA

Davidson Zhu ✉

Duke University, Durham, NC, USA

Abstract

We study the problem of selection in the context of Bayesian persuasion. We are given multiple agents with hidden values (or quality scores), to whom resources must be allocated by a welfare-maximizing decision-maker. An intermediary with knowledge of the agents' values seeks to influence the outcome of the selection by designing informative signals and providing tie-breaking policies, so that when the receiver maximizes welfare over the resulting posteriors, the expected utilities of the agents (where utility is defined as allocation times value) achieve certain fairness properties. The fairness measure we will use is majorization, which simultaneously approximately maximizes all symmetric, monotone, concave functions of the utilities. We consider the general setting where the allocation to the agents needs to respect arbitrary submodular constraints, as given by the corresponding polymatroid.

We present a signaling policy that achieves a logarithmically approximate majorized policy in this setting, assuming the receiver is a $(1 + \epsilon)$ approximate welfare maximizer. The approximation ratio is almost best possible, and that significantly outperforms generic results that only yield linear approximations. A key component of our result is a structural characterization showing that the vector of agent utilities for a given signaling policy defines the base polytope of a different polymatroid, a result that may be of independent interest. In addition, we show that an arbitrarily good additive approximation to this vector can be produced in (weakly) polynomial time via the multiplicative weights update method.

2012 ACM Subject Classification Theory of computation → Algorithmic game theory and mechanism design

Keywords and phrases Bayesian Persuasion, Fair Division, Submodular Optimization

Digital Object Identifier 10.4230/LIPIcs.FORC.2026.14

Related Version *Full Version*: <https://arxiv.org/abs/2511.08538v1> [5]

Funding *Kamesh Munagala*: NSF award IIS-2402823

Yiheng Shen: NSF award IIS-2402823

1 Introduction

The challenge of selecting fair outcomes arises in several decision-making settings, such as assembling project teams, allocating institutional funding, and recommending items or articles. Consider for example a government agency allocating research funding across different research categories and institutions. The agency has a total budget, but may impose caps on funding allocated to any single institution or research area to encourage diversity. Typically, the agency will be a welfare maximizer, allocating funding in a way that maximizes the average quality of the proposed work per dollar spent. However, the quality of proposed



© Yannan Bai, Kamesh Munagala, Yiheng Shen, and Davidson Zhu;
licensed under Creative Commons License CC-BY 4.0

7th Symposium on Foundations of Responsible Computing (FORC 2026).

Editor: Huijia (Rachel) Lin; Article No. 14; pp. 14:1–14:22

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

work is often hard to assess from the proposals, with several competing projects having comparable quality. The resulting uncertainty in assessing quality can create unintentional unfairness in allocating funds.

One way to ameliorate this problem is to carefully design the proposal mechanism to reveal the right amount of additional information about the quality of the proposed work. Indeed, revealing too much can lead to a strict winner-take-all allocation of funds, disadvantaging proposals from research areas or institutions that were only slightly inferior, while revealing too little causes many proposals, even of widely different quality, to be comparably ranked, again leading to low overall welfare, and hence unfairness.

This motivates the view of fair selection as an *information revelation* problem, an approach taken by [7, 3] in the context of selecting a single individual, for instance, in hiring or admission decisions. This information revelation problem is posed as a special case of Bayesian persuasion [23] as follows: The proposals in the above motivating example are viewed as “agents”. Each agent has a quality that is drawn from an independent prior distribution. There is an intermediary that designs the information environment, in the above case, the proposing mechanism and what the agents should reveal in it. This intermediary (called the sender) is assumed to know the exact qualities. The decision-maker (called the receiver) is the reviewing panel in the above example and only knows the prior. The intermediary constructs signals independently for each agent based on the true qualities, and sends these to the receiver. Using these signals, the receiver constructs a posterior over the qualities, subsequently choosing the winning solution. In the above case, this is the budget allocation that maximizes posterior quality per dollar allocated, while respecting the budget constraints on research areas or institutions. Ties are broken by a randomized rule specified by the intermediary.

The main research question then becomes: *How can an intermediary strategically reveal information about agents’ qualities so that a welfare-maximizing decision-maker produces a fair outcome?* Such an approach to fairness via information revelation differs from prior algorithmic work on fair selection that designed novel, often randomized, selection rules [25, 12, 30, 29, 15].

Note that the decision-maker (receiver) always acts as a welfare maximizer over its information (the posterior), while the sender guides it towards socially desirable objectives such as fairness via partial information revelation (signaling). To gain intuition for why such signaling can be beneficial, in the funding allocation example, if the quality of each proposal were exactly known through an exhaustive questionnaire, then the decision-maker could identify the absolute best candidates to fund. However, such detailed information forces the decision-maker to make a winner-take-all choice among comparable proposals based on negligible differences. In contrast, if we carefully design either the questionnaire or review process to be coarser (for instance, by letting outside expert reviewers rate proposals as “competitive” or “not competitive”), this could render many high-quality proposals indistinguishable to the decision-maker. This can let them perform randomized tie-breaking that gives each high-quality proposal a fair chance, hence ensuring fairness while not compromising overall welfare significantly. The question then becomes – how should this signal (the questionnaire or review process in this case) be designed with the fairness and welfare goals in mind?

In this work, we generalize the information design framework of [7] beyond selecting a single agent to the broader domain of selection when the allocations define a polymatroid [28, 18]. Here, the decision-maker’s goal is to choose a feasible set (possibly involving fractional or randomized allocations) that maximizes total welfare. Polymatroids capture submodularity in the allocation constraints, and hence the structure of many selection problems involving diminishing returns or diversification. In the funding allocation example, budget caps on subsets of research areas or institutions, if they are submodular, will define a polymatroid.

A special case of polymatroids is *matroid constraints*, that is, allocations that are a randomization over independent sets of a matroid. This includes choosing one candidate in hiring or admission decisions considered in [7, 3], and its generalization to selecting k candidates. Similarly, it includes partition matroids, where agents are partitioned into disjoint groups with a quota for each group. For example, in the task of hiring a specialized team of n engineers, there could be constraints that at most k_1 test engineers and at most k_2 security experts are selected. As before, the agents' quality is partially known to the decision-maker, and the agents could signal their quality via an information intermediary, with the goal of ensuring fair selection.

In the motivating example, the intermediary plays the role of a mechanism or entity that shapes the information available to the decision-maker. In settings such as hiring or grant evaluation, this corresponds to the design of questionnaires, review rubrics, or external expert ratings that determine how much detail about candidates' or proposals' true quality is revealed to the decision makers. Intermediaries also naturally arise as components such as ad exchanges or content-curation models that possess richer side information via predictive models and transmit these to the recommendation platform as described above. In all cases, the intermediary does not directly allocate resources, but influences allocation outcomes by deciding how coarsely or finely to reveal private information.

To formalize the goal of fairness, we adopt the notion of *approximate majorization* from [7]. This notion arises from economics and operations research [22, 21, 26, 13]. Given a signaling policy, we assume each agent's utility equals its quality score times its fractional allocation.¹ Each agent therefore receives an expected utility, where the expectation is over its quality and the outcome of the signaling policy. Our signaling policies will ensure the resulting vector of agents' expected utilities is approximately majorized; see Definition 4 for a formal definition. Informally, this means all symmetric, monotone, and concave fairness functions (including max-min fairness, total welfare, and Nash welfare) are simultaneously optimal to that approximation factor. In particular, such a notion is approximate on both traditional fairness notions (like max-min fairness) and the total welfare, in some sense being approximately best possible. Our goal in this paper is to design a signaling policy that achieves as small an approximation factor as possible, in a computationally efficient fashion.

As mentioned before, the key challenge in designing a fair signaling policy lies in the trade-off between too little and too much information. Our work initiates the study of fair information design for selection problems with complex allocation constraints, specifically with polymatroid constraints.

1.1 Main Results

Our main technical result is the following theorem about existence of approximately majorized solutions, which generalizes an analogous result in [7] from selecting a single agent to handle arbitrary submodular constraints:

► **Theorem 1 (Informal).** *Assuming there are n agents and each agent's quality lies in $[1, V]$, there is a $O(\log \frac{V}{\epsilon})$ approximate majorized policy when:*

- *the agents have independent distributions over quality (or value) and independent signaling policies for this quality (see Section 2 for a detailed model),*
- *the set of feasible allocations over the agents forms a polymatroid,*

¹ Our results easily generalize to the setting where agent i 's utility is vx_i , where v is a fixed multiplier and x_i is the fractional allocation to the agent.

- *the receiver is a $(1 + \epsilon)$ -approximate (in each coordinate) welfare maximizer over the polymatroid constraint given the posterior distributions over agent values, and*
- *the utility of an agent is the expectation over the signaling policy of its value times its expected fractional allocation given that value.*

The example in Appendix B shows that naive signaling policies cannot achieve the above bound even with simple polymatroidal constraints. We further note that there is a lower bound of $\Omega(\log \log V)$ on approximate majorization even for selecting a single agent [7]. This rules out an $O(1)$ approximate majorized signaling policy.

We next complement Theorem 1 by showing that an arbitrarily good approximation to the above solution can be computed in (weakly) polynomial time via an application of the multiplicative weights method.

► **Theorem 2 (Informal).** *For any $\delta > 0$, the utility vector from Theorem 1 can be approximated to an additive $O(\delta)$ in time polynomial in $1/\delta$, n , and V .*

1.2 Technical Contributions

For lack of space, complete proofs are presented in the full paper [5]. Our proof proceeds in two main stages. In Section 3, we first analyze the simpler but crucial case of “full revelation” policies, which means each agent truthfully reveals its quality. We do so to establish a novel structural property of the utility space, which we use as a building block to prove our main theorem for general signaling policies in Section 4.

Existence Result. To show the existence result in Theorem 1, in Section 3, we start with the simple setting where the signaling policy is “full revelation”. In this setting, given the revealed values of the agents, the receiver has a choice between welfare maximizing allocations and chooses a solution for each vector of revealed qualities so that in expectation over these revelations, the vector of agent utilities is as fair as possible. Our main result is Theorem 10 that shows the existence of an exactly majorized policy in this setting. The key to showing this result is Lemma 11, which provides a polyhedral characterization of persuasion with a welfare maximizing receiver:

When the underlying feasibility constraint over agent allocations defines a polymatroid and the receiver is a welfare maximizer, the set of expected utility vectors of the agents is also the base polytope of a (different) polymatroid.

Given this statement, we can leverage the existence of exactly majorized solutions for polymatroids [31, 33, 27]. To see why this statement is non-trivial, we note that in each scenario of revealed values, set of welfare maximizing allocations define a face of the polymatroidal extension, which is itself the base polytope of a polymatroid [28, 19]. However, in our setting, the utility vector of the agents is not the allocation vector, but the allocation scaled by the quality of each agent, and in general, even if the allocation vectors are drawn from a polymatroid, but are scaled by fixed quantities that depend on the index of the coordinate, the resulting vectors *do not* define a polymatroid. This makes the statement of Lemma 11 novel and non-trivial; we provide additional discussion for its subtlety in Appendix D. Our proof of Lemma 11 crucially uses the welfare maximizing behavior of the receiver and the polymatroid structure of the underlying feasibility constraint. We carefully analyze the greedy allocation rule of welfare optimization to show polymatroidal structure of the overall signaling problem. In Appendix D, we complement this result by showing

an example where the allocation set is not a polymatroid but has a 1-majorized point, but the utility vectors cannot be approximately majorized. This showcases the crucial role of submodularity in our results.

Once we establish exact majorization for full revelation policies, in Section 4, we combine this with the idea of “single-mean policies” from [7] and the existence of exactly majorized points for polymatroids [31] to show the existence of a logarithmically majorized policies for general signaling policies. Note that unlike the full revelation setting, it is no longer possible to show polymatroidal structure for the space of utilities in general signaling policies. A key ingredient here is Lemma 20 that shows the existence of a single-mean policy of a certain type that does have polymatroid structure, again using the monotonicity and submodularity properties of the underlying polymatroid. This step crucially requires the agents have independent distributions over their quality and independent signaling policies.

Computation. The ideas for showing Theorem 2 are more standard. For this, in Section 3.2, we write majorization problem as a set of linear programs over the exponentially many scenarios of revelations, with a polymatroid optimization problem (capturing receiver behavior) for each scenario. We use a majorization LP from [20]. We now use the multiplicative weights method [2] to solve this program approximately; the dual oracle becomes the expectation over scenarios, of a weighted welfare maximization problem over the base polytope of the polymatroid capturing the receiver optimization problem for that scenario. The latter can again be solved via a greedy algorithm, and the expectation can be approximated by sampling polynomially many scenarios. Our overall approach follows [11, 10], who apply similar frameworks for optimal multidimensional auctions.

Conceptual Contribution. Conceptually, Lemmas 11 and 20 allow for a direct analysis of the utility space. This is in contrast to the approach in [7], which focused on the special case of single-agent selection (a matroid special case). Their work reduced a relaxed version of the single-selection problem to a majorized network flow instance. In contrast, our proof is more direct, and provides a new structural understanding via a geometric and polymatroidal characterization of the utility space. This not only allows us to leverage existing results on finding majorized points in polymatroids [31] but also provides hope of achieving fairness and welfare guarantees for information design in other, more complex, settings.

At a higher level, the difficulty with persuasion is that the sender needs to treat the receiver’s optimization routine as a black box, which makes the optimal persuasion problem non-convex in general, even when the receiver is solving a convex optimization problem. This aspect has precluded the development of general-purpose techniques based on convex relaxations to derive structural insights into these problems. As an example, optimal auction design and pricing under persuasion requires the development of specialized techniques to handle non-convexity, and some generalizations admit strong lower bounds for this reason [1, 6]. Our main contribution is to show that a large class of submodular selection problems admits to a convex structure even in the presence of persuasion, a result that is *a priori* not obvious.

1.3 Other Related Work

Bayesian Persuasion. *Information design* is a framework for understanding how a sender can influence a receiver’s actions by strategically revealing information [9, 16]. Within this broad area, our work falls in the setting of Bayesian persuasion [23], where the receiver performs Bayesian updates based on the sender’s signals. This problem has been widely

studied in various contexts in computer science and economics [8, 6, 34, 4, 14, 32]. We note that computationally efficient algorithms exist for arbitrary objectives in persuasion, notably the FPTAS of [17]. However, majorization requires the simultaneous near-optimality of an entire family of fairness functions, and here, even showing existence is non-trivial. Making progress therefore requires the development of new structural insights into the problems that deviate significantly from prior literature. As mentioned before, our work derives novel convexity characterizations for a large class of persuasion problems, allowing us to argue strong fairness properties.

Our model builds on work by [3], who consider selfish agents who independently construct their own signaling policies to persuade a receiver to allocate to them. In contrast with their work that focuses on allocating to one agent, we consider general allocation polymatroids, and focus on fairness in a centralized setting with a common sender, akin to [7].

Majorization in Optimization. Majorization was introduced in the seminal works of [24, 22]. It provides a strong framework for fairness that is equivalent to maximizing all symmetric and concave welfare functions. The work of [21, 20] defined an approximate version suitable for resource allocation, hence applying it to approximation algorithms. More classical work has found connections between majorization and specific combinatorial structures, most notably the exact majorization of flows in single-source multi-sink networks [33, 27], and of polymatroids in general [31]. Our work contributes to this literature by establishing a new structural connection between majorization in Bayesian persuasion and the geometry of polymatroids. The resulting approximation ratios are a significant improvement over generic approximation bounds that follow from [20], that can depend linearly on problem parameters.

2 Preliminaries

2.1 Signaling Policies

There is a set E of n agents. We call the decision-maker the receiver. The value v_i of each agent i is drawn independently from the distribution D_i . The decision-maker knows the distributions $\{D_i\}_{i=1}^n$, but does not know the realized values $\{v_i\}_{i=1}^n$. We assume $v_i \in [1, V]$ for all i .

After the values $\{v_i\}_{i=1}^n$ are realized, an intermediary (or sender) uses these values to send signals $\{\sigma_i\}_{i=1}^n$ to the receiver via a *signaling policy*. A signaling policy ω comprises the *mapping rule* and the *selection rule*. The signal for each agent is independently constructed from the other agents.

Mapping Rule. A mapping rule is a collection of signals for each agent $\{\Gamma_i\}_{i=1}^n$ together with a function that maps the value v_i of an agent i to a distribution g_{iv_i} over signals in Γ_i . When the sender sees the realized values $\{v_i\}_{i=1}^n$, they compute the corresponding signal distributions $\{g_{iv_i}\}_{i=1}^n$ by the mapping rule. They then generate the realized signals $\{\sigma_i\}_{i=1}^n$ by drawing $\sigma_i \sim g_{iv_i}$ independently for each agent, and the receiver sees $\{\sigma_i\}_{i=1}^n$.

After receiving the set of signals, the receiver computes the posterior distributions $\{D_i(\sigma_i)\}_{i=1}^n$ over agent values using Bayes' rule. Let $\mu_i = \mathbb{E}[D_i(\sigma_i)]$ denote the posterior mean of agent i . A set of posterior means is said to be *Bayes plausible* if it corresponds to a valid signaling policy. Under Bayes plausibility, the expectation of the posterior mean over the signals is equal to the prior mean.

Allocation Constraints. There is a polymatroid constraint (E, f) on the set of possible allocations to the agents.

► **Definition 3** (Submodularity and Polymatroids). *Let E be a finite set and f a non-negative, monotone, submodular, function from the power set 2^E to $\mathbb{R}_+$, which satisfies $f(\emptyset) = 0$;*

$$f(A) \leq f(B) \text{ for } A \subseteq B \subseteq E, \text{ and}$$

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \text{ for } A, B \subseteq E.$$

Then, the pair (E, f) is called a polymatroid, where E is called the ground set and f the rank function of the polymatroid. A polymatroid defines a polytope $\mathcal{P}(f) \subset \mathbb{R}_+^E$ by

$$\mathcal{P}(f) = \{\mathbf{x} : \mathbf{x}(A) \leq f(A) \text{ for all } A \subseteq E\}.$$

This polytope is called the independence polytope of the polymatroid. When there is no ambiguity, we also refer to the independence polytope as the polymatroid.

The base polytope of a polymatroid (or the corresponding submodular function) is the following:

$$\mathcal{B}(f) = \{\mathbf{x} \in \mathcal{P}(f) : \mathbf{x}(E) = f(E)\}.$$

Selection Rule. We assume that the receiver is a utilitarian welfare maximizer, so that it maximizes the sum of the posterior utilities of the agents subject to the polymatroid constraint (E, f) . In other words, the receiver constructs a welfare-maximizing allocation $\mathbf{x} \in \mathcal{P}(f)$ that maximizes $\sum_{i=1}^n \mu_i x_i$. The set of welfare-optimal allocations $\mathbf{x}$ defines a face of the polymatroid (E, f) , and the receiver chooses an allocation from this face to satisfy the sender's fairness objective (maximizing the majorization of the utility vector, see Definition 4 below). This choice is termed the *selection rule*. The assumption of sender-preferred tie-breaking is standard in the Bayesian persuasion literature [23], and is consistent with a receiver who is intrinsically motivated by fairness but constrained to be welfare-optimal. We assume agent i obtains utility $\mu_i x_i$.

A signaling policy Ω is a distribution over independent signaling policies ω . Before the process starts, the sender draws $\omega \sim \Omega$ and implements ω . The signaling policy is known to the receiver. Since $v_i \sim D_i$, this yields a expected utility $U_i(\Omega)$ for the agent, where the expectation is over D_i , the distributions of other agents' values, and the distribution over signaling policies in Ω .

2.2 Fairness and Majorization

The goal of the sender is to design a signaling policy Ω that is fair. We capture this as designing Ω such that the vector $\{U_i(\Omega)\}_{i=1}^n$ is α -majorized over the set of all signaling policies, for the smallest possible value α . The selection rule of $\mathbf{x}$ among the receiver's welfare-maximizing allocations will be influenced by this fairness goal.

► **Definition 4** (α -Majorization, [20, 7]). *For $\alpha \geq 1$, a signaling policy Ω is called α -majorized if for any $k \in \{1, 2, \dots, n\}$ and any signaling policy Ω' , the sum of the k smallest utilities in $\{U_i(\Omega)\}_{i=1}^n$ is at least $1/\alpha$ times the sum of the k smallest utilities in $\{U_i(\Omega')\}_{i=1}^n$.*

The following result shows that approximate majorization is equivalent to simultaneously approximating all symmetric and concave welfare functions.

► **Proposition 5** ([20]). *The signaling policy Ω is α -majorized if and only if for every symmetric, non-decreasing, and concave function $h: \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$ and any other signaling policy Ω' ,*

$$h(U(\Omega)) \geq \frac{1}{\alpha} \cdot h(U(\Omega')).$$

To see why active signal design is required, consider the trade-off illustrated in the example in Appendix B.

2.3 Properties of Polymatroids

Given a polymatroid $\mathcal{P}(f)$, the classic greedy algorithm works as follows:

1. Order the indices of $[n]$ according to a permutation (ordering) π .
2. For $k = 1, \dots, n$, set $x_{\pi(k)} = f(\{\pi(1), \dots, \pi(k)\}) - f(\{\pi(1), \dots, \pi(k-1)\})$.

We have the following well-known lemmas:

► **Lemma 6** ([28], Chapter 44). *Given a vector $\vec{v} > \vec{0}$, the function $\vec{v} \cdot \vec{x}$ is maximized over $\mathcal{P}(f)$ by the greedy algorithm that uses an ordering π such that $v_{\pi(1)} \geq v_{\pi(2)} \dots \geq v_{\pi(n)}$. Further, any vertex² of the optimal face of $\mathcal{P}(f)$ corresponds to some permutation π satisfying $v_{\pi(1)} \geq v_{\pi(2)} \dots \geq v_{\pi(n)}$.*

► **Lemma 7** ([18], Chapter 3). *The set of vertices of $\mathcal{B}(f)$ coincides with the set of vectors $\vec{x}$ obtained by running the greedy algorithm for all possible orderings π .*

In the classic lemma below, the result for the polytopes $\mathcal{P}$ is from [28], while the result for base polytopes uses the above characterization of its vertices – it is easy to write each vertex of $\mathcal{B}(f_1 + f_2)$ as the sum of the corresponding vertices of $\mathcal{B}(f_1) + \mathcal{B}(f_2)$.

► **Lemma 8** ([28]). *The following statements hold for independence and base polyhedra of non-negative, monotone, submodular functions:*

- $\mathcal{P}(f_1) + \mathcal{P}(f_2) = \mathcal{P}(f_1 + f_2)$ and $\mathcal{B}(f_1) + \mathcal{B}(f_2) = \mathcal{B}(f_1 + f_2)$;
- For any $\alpha > 0$, we have $\mathcal{P}(\alpha f) = \alpha \mathcal{P}(f)$ and $\mathcal{B}(\alpha f) = \alpha \mathcal{B}(f)$.

The following lemma captures the relation between polymatroids and majorization.

► **Lemma 9** ([31]). *Any polymatroid (E, ρ) has a 1-majorized element that lies in $\mathcal{B}(\rho)$.*

3 Full Revelation Signaling Policies

We first show that within the restricted class of *Full Revelation Policies*, there exists a solution that 1-majorizes all other solutions in that same class. We then show an approximation algorithm to compute it. This will form the basis of the proof of approximate majorization (and the associated computational result) for general policies in Section 4.

The mapping rule of a full revelation policy is directly sending the realized value to the receiver, and the signaling policy involves designing the selection rule for the receiver. Assume that there are only finitely many possible values for all agents, denoted by $v_1 > v_2 > \dots > v_k > 0$. Let v_{a_i} be the value sent by agent i . In this setting, note that if v_{a_i} is the revealed value of agent i , then the posterior mean is simply $\mu_i = v_{a_i}$, and the receiver selects an $\mathbf{x} \in \mathcal{P}(f)$ that maximizes $\sum_i v_{a_i} x_i$, breaking ties in favor of the majorization objective. This yields expected utility vector $\{U_i(\Omega)\}_{i=1}^n$ for the agents.

² This is shown for the extended polymatroid in [28]. Note that for a strictly positive $\vec{v}$, the optimum face will belong to the extended polymatroid.

3.1 Existence of a 1-Majorized Solution

We will show the following theorem.

► **Theorem 10** (Existence of a 1-majorized policy). *Assume the allocation constraints define a polymatroid (E, f) . Let $\mathcal{U}_{FR}$ be the set of expected utility vectors achievable by full-revelation signaling policies (where the mapping rule is fixed to full revelation and the selection rule varies). Then, the set $\mathcal{U}_{FR}$ contains a vector that 1-majorizes every other vector in $\mathcal{U}_{FR}$.*

Fix a set of realized values of the agents. We will show that the vector of utilities of the agents forms the base polytope of a different polymatroid.

► **Lemma 11.** *Let $\mathcal{P} = \mathcal{P}(f)$ be a polymatroid on a ground set E , defined by a submodular rank function f . Let $\vec{v}$ be a strictly positive vector of agent values. Let $\mathcal{X}^*$ be the face of optimal allocations in $\mathcal{P}$ that maximize the welfare function $v \cdot x$. Let the corresponding set of utility vectors be $\mathcal{U} = \{(v_i x_i)_{i \in E} \mid x \in \mathcal{X}^*\}$. Then, the set $\mathcal{U}$ is the base polytope of a submodular function.*

The proof is deferred to Appendix A.

Proof of Theorem 10. Note that $\{U_i(\Omega)\}_{i=1}^n$ is the Minkowski sum of the vectors $\{\Pr[\sigma] \cdot U_i(\sigma)\}$, where σ are scenarios of realized values. By Lemma 8, scaling a base polytope by a constant is also a base polytope, and so is taking Minkowski sums. Combining with Lemma 11, this means the set $\{U_i(\Omega)\}_{i=1}^n$ also defines the base polytope of a polymatroid, and by Lemma 9, this base polytope contains a point that 1-majorizes all other points within the polytope. ◀

Indeed, by combining Lemma 8 with Lemma 11, the set of vectors of expected utilities $\vec{U}$ of the agents for feasible signaling policies coincides with the base polytope of the following polymatroid $\mathcal{R}$. Here, $g(S; \vec{v})$ is the function $g(S)$ from Lemma 11 when the realized value vector is $\vec{v}$. The expectation below is over the realized value vector.

$$\mathcal{R} = \left\{ \vec{y} \geq 0 \mid \sum_{i \in S} y_i \leq \mathbb{E}_{\vec{v}} [g(S; \vec{v})] \quad \forall S \subseteq [n] \right\}. \quad (1)$$

Lemma 11 implies $\mathbb{E}[g(S; v)]$ is submodular, so that the above set of constraints define a polymatroid, and has a 1-majorized point.

► **Remark 12.** We note that the proof of Lemma 11 is quite delicate. In Appendix D, we present two examples to support this.

3.2 Polynomial Time Approximation Scheme

We will next show a polynomial time additive approximation to compute the 1-majorized point.

► **Theorem 13.** *In the full information revelation setting, we can compute a policy that approximates the 1-majorized vector of utilities to an additive $O(\delta)$ in time $\text{poly}(n, V, 1/\delta)$.*

The above theorem also shows that the approximation ratio can easily be made multiplicative $(1 + \delta)$ if the running time is $\text{poly}(n, V/\delta, \text{OPT}_n/\text{OPT}_1)$, where OPT_1 is the max-min fair utility value and OPT_n is the social welfare.

In the rest of the section, we provide a proof sketch of Theorem 13. Assume that all values are normalized so that the smallest value is 1 and the largest value is V . The algorithm is an application of the multiplicative weights update framework. We use this to compute the

14:10 Fair Multi-Agent Persuasion with Submodular Constraints

maximum sum of each of the smallest k utilities by binary sum, and then use the framework again with these values to compute a sequence of policies that can be computed efficiently by sampling, whose average approximates the majorized solution to an additive δ . Since the details of our approach are very similar to that in [10], we only present a sketch and omit the details.

Given a vector $\vec{x} = (x_1, \dots, x_n)$, let the i -th smallest element of x be $x_{(i)}$. We define $Q_j(\vec{x}) = \sum_{i=1}^j x_{(i)}$. Recall that $\vec{x}$ is majorized by $\vec{y}$ or $\vec{y}$ α -majorizes $\vec{x}$ if $\alpha \cdot Q_j(\vec{y}) \geq Q_j(\vec{x})$.

Let $\vec{v} = \{v_{a_i}\}_{i=1}^n$ be the realized values of the agents. Using a result in [20], for every $1 \leq j \leq n$, the program below finds $\max\{Q_j(\{U_i(\Omega)\}) : \Omega \text{ feasible policy}\}$:

$$\begin{aligned} & \text{Maximize } \left(\sum_{i=1}^n U'_i \right) - (n-j)M \text{ subject to:} \\ & U_i \leq \mathbb{E}_{\vec{v}} \left[\sum_{i=1}^m v_{a_i} x_{i\vec{v}} \right], \text{ for all } i \\ & \{v_{a_i} x_{i\vec{v}}\} \in \mathcal{F}(\vec{v}), \text{ for all } \vec{v} \\ & U'_i \leq \min\{U_i, M\}, \text{ for all } i. \end{aligned}$$

Here, $\mathcal{F}(\vec{v})$ is the base polytope $\mathcal{B}(g)$ from the proof of Lemma 11 when the realized value vector is $\vec{v}$. Let OPT_j^* denote the optimal objective to the above program.

► **Lemma 14** (Lemma 3.1 in [20]). *The linear program above finds $\max\{Q_j(\{U_i(\Omega)\}) : \Omega \text{ feasible policy}\}$.*

We now follow the framework in [10] and use the Multiplicative Weights method to decide feasibility of the first constraint subject to all the others.

For a fixed guess objective value OPT_j (which we can find the optimal value of via binary search), we rewrite the above LP as a feasibility problem for the objective being at least OPT_j . It suffices to solve the corresponding oracle problem with nonnegative dual multipliers $\{\lambda_i\}$:

$$\begin{aligned} & \text{Maximize } - \sum_i \lambda_i U_i + \mathbb{E}_{\vec{v}} \left[\sum_i \lambda_i v_{a_i} x_{i\vec{v}} \right] \\ & \{v_{a_i} x_{i\vec{v}}\} \in \mathcal{F}(\vec{v}), \text{ for all } \vec{v} \\ & U'_i \leq \min\{U_i, M\}, \text{ for all } i \\ & \left(\sum_{i=1}^n U'_i \right) - (n-j)M \geq \text{OPT}_j. \end{aligned}$$

This optimization problem decouples into two separate optimization programs. Minimizing $\sum_i \lambda_i U_i$ subject to all but the first constraint is a linear program and can be solved in polynomial time. Finding the maximum of $\mathbb{E}_{\vec{v}} [\sum_i \lambda_i v_{a_i} x_{i\vec{v}}]$ splits into finding the maximum of $\sum_i \lambda_i v_{a_i} x_{i\vec{v}}$ subject to $\{v_{a_i} x_{i\vec{v}}\} \in \mathcal{F}(\vec{v})$ for each $\vec{v}$. The work of [10] shows that for parameter $\delta > 0$, we can find a solution to the original LP with value at least OPT_j^* satisfying

$$U_i \leq \mathbb{E}_{\vec{v}} \left[\sum_{i=1}^m v_{a_i} x_{i\vec{v}} \right] + \frac{\delta}{n}$$

in time $\text{poly}(n, V, 1/\delta)$. The details beyond this point are similar to [10] and we defer them to the full paper [5].

4 General Signaling Policies and Approximate Majorization

We now build on the results in Section 3 to show Theorems 1 and 2, the existence of approximate majorized policies, and associated computational result, for general policies. To appreciate the technical challenge, the mapping scheme in full revelation policies is fixed, so that we only need to focus on designing the selection policy (or allocation rule) of the receiver. This makes the overall problem have polymatroid structure if it has that structure for a fixed scenario. However, for general signaling policies, there is a dependence between the mapping rule in the signaling policy and the allocation made by the receiver. Since both the mapping rule and selection rule are not fixed anymore, the overall signaling problem may not have polymatroid structure.

To extend our result to approximate majorization of general policies, we adopt the approach of randomized single mean projections introduced in [7] for selecting a single agent. For these policies, it was shown that there is a fixed set of mappings, termed maximal mappings, that can be pre-computed and are optimal (in terms of majorization) within this class. These mappings allowed them to approximate any mapping policy by an analog of full revelation policies. We follow this outline; however, we need a different set of technical arguments to show that this class of policies suffice. The main novelty in our case, beyond extending Lemma 11 to single-mean policies, is the proof of Lemma 20, which carefully uses submodularity to show that it suffices to consider maximal mappings.

4.1 Single Mean Projections

The definitions in this section mirrors that in [7]. We briefly review the definitions for completeness. Recall that the values v_i of the agents are supported on $[1, V]$. Intuitively, single mean projection partitions the value range $[1, V]$ into a sequence of buckets, and only counts utility from one bucket. Given small $\varepsilon > 0$, let $\eta = 1 + \varepsilon$. Assume V is a power of η . Divide $[1, V]$ into buckets $I_1 = [1, \eta), I_2 = [\eta, \eta^2), \dots, I_k = [V/\eta, V)$. Let $K = O\left(\frac{\log V}{\eta}\right)$ denote the number of buckets. We will use these buckets $\{I_k\}_{k=1}^K$ to partition the range of posterior means, where each bucket I_k is associated with a canonical mean value m_k . We will construct signaling policies that choose a bucket at random and focus on the case where the posterior mean lies within this bucket.

Approximate Welfare-maximizing Receiver. As discussed in Section 1, a key element of our approach to generalizing to arbitrary signals is to model the receiver not as a perfect optimizer over the exact posterior means, but as an approximate one who acts on canonical values.

This models a receiver who is a $(1 + \epsilon)$ approximate welfare maximizer in each dimension, acting on canonical values rather than exact posterior means. From now on, we will ignore the $(1 + \epsilon)$ factor, and assume the utility of an agent is computed using the canonical posterior mean values.

Single-Mean Projections. Suppose the receiver is an approximate welfare maximizer as defined above. We define a single-mean policy for bucket I_k as follows:

► **Definition 15** (Single-mean Policies and Fake Utilities). *Consider any signaling policy specified by a mapping rule and a selection rule. For any bucket I_k , the corresponding single-mean policy restricted to that bucket accounts for the utility of any agent as follows.*

14:12 Fair Multi-Agent Persuasion with Submodular Constraints

The fake utility of an agent i , denoted by $\hat{U}_{i,k}$, is measures as the utility of agent i when the posterior mean lies in I_k , else zero. Formally, if $\vec{x}$ denotes the allocation, then

$$\hat{U}_{i,k}(\vec{x}) = \begin{cases} m_k \cdot x_i, & \text{if } \mu_i \in I_k; \\ 0, & \text{if } \mu_i \notin I_k. \end{cases}$$

Note that the only difference in a single-mean policy and a regular policy is the utility accounting as a fake utility. This fake utility serves as an underestimation of the true utility seen by the agent in the policy. We now define a randomized single-mean policy as follows:

► **Definition 16** (Randomized Single-mean Policies). A randomized single-mean policy is constructed as follows: The mapping rule Δ is a collection of K mapping rules $\Delta_1, \dots, \Delta_K$ and associated selection rules, yielding a collection of signaling policies $\Omega_1, \dots, \Omega_K$. The utility of policy Ω_k is measured using the fake utility restricted to bucket I_k . The overall policy chooses one of the K buckets uniformly at random, and uses the corresponding signaling policy Ω_k .

Note that the expected fake utility of agent i in the above randomized single-mean policy is

$$\hat{U}_i(\Omega) = \frac{1}{K} \sum_{k=1}^K \hat{U}_{i,k}(\Omega_k).$$

Note that given any signaling policy Ω , there is a randomized single-mean policy Ω_{rsm} obtained by picking a bucket uniformly at random and using the fake utility restricted to that bucket. In this policy, we have $\hat{U}_i(\Omega_{\text{rsm}}) = \frac{U_i(\Omega)}{K}$ for all agents i , where $\hat{U}_i(\Omega_{\text{rsm}})$ is the expected fake utility of agent i in the randomized single mean policy. Note that we have accounted for $U_i(\Omega)$ by rounding each posterior mean to its canonical value, and ignored the $(1 + \epsilon)$ factor loss in this process.

We now consider fixing the mapping rule Δ used by the randomized single-mean policies, but do not fix the selection rule used by the receiver. In other words, for every bucket I_k , we specify the mapping rule Δ_k of values to signals. Note that this bucket is chosen with probability $1/K$ in the rule Δ . The following is analogous to Lemma 11.

► **Lemma 17.** Fix an active bucket I_k , and the mapping rule Δ_k of the corresponding single mean policy. Fix the vector of posterior means μ of the agents. Let $\hat{g}_k(S; \mu)$ denote the maximum possible sum of the fake utilities of agents in a set S , where the maximization is over the selection rule of the receiver. In other words,

$$\hat{g}_k(S; \mu) = \max \sum_{i \in S} \hat{U}_{i,k}$$

where $\hat{U}_{i,k}$ is as defined in Definition 15. Then, the function $\hat{g}_k(S; \mu)$ is monotone and submodular.

Taking the Minkowski sum over the random choice of bucket I_k and over the realized posterior means, and using Lemma 8, we obtain the following.

► **Corollary 18.** Given a randomized single mean mapping rule $\Delta = \{\Delta_k\}_{k=1}^K$, let the expected fake utility be $\hat{U}_i = \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{\mu \sim \Delta_k} [\hat{U}_{i,k}]$. Let $\mathcal{U}$ denote set of vectors $\{\hat{U}_i\}_{i=1}^n$ obtained by varying the selection rule of the receiver. Then, $\mathcal{U}$ is contained in the following polymatroid $\hat{\mathcal{R}}(\Delta)$, and contains its base polytope, where $\hat{\mathcal{R}}(\Delta)$ is defined as:

$$\hat{\mathcal{R}}(\Delta) = \left\{ \vec{y} \geq 0 \mid \sum_{i \in S} y_i \leq \frac{1}{K} \sum_{k=1}^K \mathbb{E}_{\mu \sim \Delta_k} [\hat{g}_k(S; \mu)] \forall S \subseteq [n] \right\}. \quad (2)$$

Here, the expectation in the RHS is over the vector $\vec{\mu}$ of posterior means produced by the mapping Δ_k , and the functions $\hat{g}_k$ are as defined in Lemma 17.

4.2 Maximal Single-mean Mappings

The issue with extending the above lemma to a proof analogous to Theorem 10 is that the mapping rule Δ is now a variable (in addition to the selection rule that depends on Δ). This wasn't an issue in the proof of Theorem 10, where the mapping rule was fixed and the set of utilities obtained by varying the selection rule defines the base of a polymatroid. In contrast, though $\hat{\mathcal{R}}(\Delta)$ as defined above is a polymatroid, the union of such polymatroids over Δ need not have nice structure.

We now proceed as in [7] and show that the optimal signaling policy for single-mean policies is fixed and independent of the allocation. This will allow us to argue polymatroidal structure, and show that the space of randomized single mean policies has a 1-majorized solution. Since the true utility is within a factor of K of the fake utilities used by such policies, this directly implies a K -majorized policy for general signaling policies, and we will show that in Theorem 22.

Towards this end, we define a *maximal mapping* analogous to [7].

► **Definition 19** (Maximal Mapping). *For an interval I_k , a maximal mapping is a mapping rule ω from agent values to signals $\{\sigma\}$ such that $\Pr_{\sigma \sim \omega}[\mu_i(\sigma) \in I_k]$ is maximized for each agent i .*

Note that for each agent i , the maximal mapping to a given interval I_k is the solution to a linear program [7]. This mapping is fixed and decoupled from the allocation rule of the receiver. It can also be computed separately for each agent. The set of maximal mappings, one for each I_k , yields the mapping rule of a randomized single-mean policy, by choosing one of the buckets uniformly at random. We call this mapping rule $\Delta_{\max}$.

We now present the key structural lemma to show that randomized single mean policies can switch to using maximal mappings without reducing the maximum expected utility of any set of agents. In the lemma below, the notation $\hat{g}_k(S; \mu)$ is as defined in the proof of Lemma 17. Further, by the notation $\mu \sim \Delta_k$, we mean a posterior mean vector that results from the execution of the mapping rule Δ_k .

► **Lemma 20** (Structure Lemma). *Consider a fixed active bucket I_k . Let $\Delta_k = (\sigma_1, \dots, \sigma_n)$ be an arbitrary mapping rule where each agent's mapping σ_i is chosen independently. Consider the mapping rule $\Delta_k^{\max} = (\sigma'_1, \dots, \sigma'_n)$, where each σ'_i is a maximal mapping for agent i with respect to the bucket I_k . Then, for any set S of agents, we have:*

$$\mathbb{E}_{\mu \sim \Delta_k^{\max}}[\hat{g}_k(S; \mu)] \geq \mathbb{E}_{\mu \sim \Delta_k}[\hat{g}_k(S; \mu)].$$

The above lemma implies the following corollary:

► **Corollary 21.** *Consider any randomized single mean policy Ω and let $\vec{U}$ denote the vector of expected (fake) utilities of the agents in this policy. Then $\vec{U} \in \hat{\mathcal{R}}(\Delta_{\max})$, where $\Delta_{\max}$ is the mapping obtained by choosing one of the K buckets uniformly at random and using the corresponding maximal mapping for each agent.*

Proof. The vector $\vec{U}$ belongs to the polymatroid $\hat{\mathcal{R}}(\Delta)$ in Equation (2), where the mapping Δ corresponds to the policy Ω . By Lemma 20, if we use the maximal mapping in each bucket instead, the RHS of the constraints in Equation (2) do not decrease. This means $\vec{U} \in \hat{\mathcal{R}}(\Delta_{\max})$. ◀

4.3 Main Result: Proof of Theorems 1 and 2

We now combine the structural results from the preceding sections to formally state and prove our main theorem, which establishes the existence of a computationally efficient signaling policy with a logarithmic approximation guarantee for majorization. We again note that this result is complemented by a lower bound from [7] that rules out an $o(\log \log V)$ -majorized policy even for selecting one agent.

► **Theorem 22.** *Consider the signaling problem with a polymatroid constraint $\mathcal{P}(f)$, where agents have independent quality distributions supported on $[1, V]$. Assume the receiver is a $(1 + \varepsilon)$ -approximate welfare maximizer who acts on canonical posterior means derived from a partition of $[1, V]$ into $K = O((\log V)/\varepsilon)$ buckets. Then, there exists a signaling policy Ω that is $O((\log V)/\varepsilon)$ -majorized over the set of all possible independent signaling policies. Furthermore, a policy that yields an additive $O(\delta)$ approximation to this utility vector can be computed in time polynomial in $\frac{nV}{\varepsilon\delta}$.*

Proof. The proof consists of two parts. First, we prove the existence of a policy with the stated approximation guarantee by relating any optimal policy to the randomized single mean policies described above. Second, we argue that this policy can be computed in polynomial time using the multiplicative weights update framework from Section 3.

Proof of Theorem 1. We now show the existence result. Let Ω^* be any signaling policy. Let $U(\Omega^*)$ be the vector of true expected utilities for this optimal policy. As discussed before, consider a randomized single mean policy, Ω_{rsm} , which is constructed from Ω^* . This policy works by first choosing a bucket $k \in \{1, \dots, K\}$ uniformly at random and creating a fake utility function that only grants the utility an agent would have received from that specific bucket in the original policy Ω^* . The expected utility for agent i under this constructed policy is $U(\Omega_{\text{rsm}}) = U(\Omega^*)/K$.

By Corollary 21, the utility vector $\{U(\Omega_{\text{rsm}})\}_{i=1}^n \in \hat{\mathcal{R}}(\Delta_{\text{max}})$, where Δ_{max} is the mapping rule that first chooses a bucket k uniformly at random, and then implements the maximal mapping rule for that bucket. Consider the class of signaling policies $\mathcal{C}_{\text{max}}$ that use Δ_{max} as their mapping rule. In such policies, the mapping rule is now decoupled from the selection rule since the mapping Δ_{max} can be pre-computed. By Corollary 18, the set of expected utility vectors achievable by policies in $\mathcal{C}_{\text{max}}$ lies within the polymatroid $\hat{\mathcal{R}}(\Delta_{\text{max}})$ and contains its base polytope, and this set is non-empty. By Lemma 9, this base polytope has a signaling policy, call it Ω_{maj} , which is 1-majorized over all policies in $\mathcal{C}_{\text{max}}$, and hence over all vectors in $\hat{\mathcal{R}}(\Delta_{\text{max}})$. In particular, this means $U(\Omega_{\text{maj}})$ majorizes $U(\Omega_{\text{rsm}})$, which is at least $U(\Omega^*)/K$. This means the utility vector $U(\Omega_{\text{maj}})$ majorizes $U(\Omega^*)/K$. This proves the existence of a policy that is $O\left(\frac{\log V}{\varepsilon}\right)$ -majorized, since $K = O\left(\frac{\log V}{\varepsilon}\right)$.

Proof of Theorem 2. We next sketch the computational result. We use the multiplicative weights approach from Section 3.2, where we use the posterior (bucketed) mean vector μ' found by the maximal mapping instead of the value vector. The core requirement for such methods to be efficient is the existence of a polynomial-time oracle for maximizing any linear function over the following polymatroid. Given μ' , the polymatroid has rank function $\hat{g}_k(S; \mu')$. The dual oracle must solve $\max_u w \cdot u$ over this polymatroid for a given weight vector w . This can be solved by the polymatroid greedy algorithm, which requires oracle access to the rank function $\hat{g}_k(S; \mu')$, followed by sampling scenarios μ' . By combining this with a binary search over the optimal utility values (the prefix sums Q_j), we obtain a polynomial-time additive approximation scheme for computing the desired logarithmically-approximate majorized policy. We omit the details as they are similar to the proof in Section 3.2. This concludes the proof of Theorem 22. ◀

5 Extensions and Open Questions

Our main contribution is a structural characterization of the utility space in Bayesian persuasion with polymatroid constraints, showing it forms a base polytope of a different polymatroid. This enabled a direct geometric approach to construct a logarithmically-approximate majorized signaling policy. This result highlights a new connection between the geometry of information design and the combinatorial structure of submodular optimization, with potential applications to other information design problems.

Our techniques easily extend to the setting where the utility of an agent is a fixed multiplier of its allocation, rather than allocation multiplied by the quality (or value). The former case is simpler, since the utility vector now coincides with the allocation vector (appropriately scaled). For a welfare maximizing receiver, the resulting set of utility vectors is trivially a face of $\mathcal{P}(f)$ and is hence the base polytope of a polymatroid. This observation extends the results in the paper to show the same approximation factor for majorization.

Our work leaves several questions open. One open question is to extend our results to the case where the intermediary can correlate the signals between agents. Another question is to understand the combinatorics of the induced fairness polyhedron. Our proof constructs its rank function g , but we do not study its interpretation. For instance, if the original constraint is a randomization over independent sets of a matroid, then how is the induced base polytope related to the original matroid? Next, can we design efficient algorithms to find a specific point that maximizes, for instance, the Nash welfare or max-min fairness? Our results imply a logarithmic approximation in polynomial time, but it is likely these problems admit to a FPTAS.

At a higher level, it would be interesting to explore other models of allocation. For instance, what if information revelation has a cost, so that, say, the sender's signals are constrained to focus on a few agents? Similarly, what if the agents arrive one at a time, with both the sender and the receiver knowing their priors upfront, while the receiver has to make irrevocable allocations to each arriving agent based on its signal? Finally, can we apply fair persuasion to settings where the receiver is solving a stochastic optimization problem, where for instance, performing two-stage optimization to design a network over the agents, or running a prophet pricing algorithm over the agents [32]. These questions offer a rich domain for structural and algorithmic inquiry.

Acknowledgment of AI use. We have used an LLM (Gemini 2.5 Pro) in this work for identifying and summarizing prior work and paraphrasing and polishing the text. All AI-generated content was verified by the authors, who take full responsibility for its correctness.

References

- 1 Reza Alijani, Siddhartha Banerjee, Kamesh Munagala, and Kangning Wang. The limits of an information intermediary in auction design. In *Proceedings of the 23rd ACM Conference on Economics and Computation (EC)*, pages 849–868. ACM, 2022. doi:10.1145/3490486.3538370.
- 2 Sanjeev Arora, Elad Hazan, and Satyen Kale. The Multiplicative Weights Update Method: a Meta Algorithm and Applications. *Theory of Computing*, 2012. doi:10.4086/toc.2012.v008a006.
- 3 Pak Hung Au and Keiichi Kawai. Competitive information disclosure by multiple senders. *Games and Economic Behavior*, 119:56–78, 2020. doi:10.1016/J.GEB.2019.10.002.
- 4 Yakov Babichenko, Inbal Talgam-Cohen, Haifeng Xu, and Konstantin Zabarnyi. Regret-minimizing bayesian persuasion. In *Proceedings of the 22nd ACM Conference on Economics and Computation (EC)*, page 128. ACM, 2021. doi:10.1145/3465456.3467574.

- 5 Yannan Bai, Kamesh Munagala, Yiheng Shen, and Davidson Zhu. Fair multi-agent persuasion with submodular constraints, 2025. doi:10.48550/arXiv.2511.08538.
- 6 Siddhartha Banerjee, Kamesh Munagala, Yiheng Shen, and Kangning Wang. Fair price discrimination. In *Proceedings of the 2024 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2679–2703. SIAM, 2024. doi:10.1137/1.9781611977912.96.
- 7 Siddhartha Banerjee, Kamesh Munagala, Yiheng Shen, and Kangning Wang. Majorized bayesian persuasion and fair selection. In Yossi Azar and Debmalaya Panigrahi, editors, *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2025, New Orleans, LA, USA, January 12-15, 2025*, pages 1837–1856. SIAM, 2025. doi:10.1137/1.9781611978322.57.
- 8 Dirk Bergemann, Benjamin Brooks, and Stephen Morris. The limits of price discrimination. *American Economic Review*, 105(3):921–957, 2015.
- 9 Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.
- 10 Anand Bhalgat, Sreenivas Gollapudi, and Kamesh Munagala. Optimal Auctions via the Multiplicative Weight Method, 2013. doi:10.48550/arXiv.1211.1699.
- 11 Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Optimal Multi-dimensional Mechanism Design: Reducing Revenue to Welfare Maximization. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 130–139, New Brunswick, NJ, USA, 2012. IEEE. doi:10.1109/FOCS.2012.88.
- 12 L. Elisa Celis, Anay Mehrotra, and Nisheeth K. Vishnoi. Interventions for ranking in the presence of implicit bias. In *Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency (FAT*)*, pages 369–380. ACM, 2020. doi:10.1145/3351095.3372858.
- 13 Deeparnab Chakrabarty and Chaitanya Swamy. Approximation algorithms for minimum norm and ordered optimization problems. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 126–137. ACM, 2019. doi:10.1145/3313276.3316322.
- 14 Archishman Chakraborty and Rick Harbaugh. Persuasive puffery. *Marketing Science*, 33(3):382–400, 2014. doi:10.1287/MKSC.2013.0826.
- 15 Siddhartha Devic, Aleksandra Korolova, David Kempe, and Vatsal Sharan. Stability and multigroup fairness in ranking with uncertain predictions. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*. PMLR, 2024.
- 16 Shaddin Dughmi. Algorithmic information structure design: a survey. *SIGecom Exchanges*, 15(2):2–24, 2017. doi:10.1145/3055589.3055591.
- 17 Shaddin Dughmi and Haifeng Xu. Algorithmic bayesian persuasion. In Daniel Wicks and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 412–425. ACM, 2016. doi:10.1145/2897518.2897583.
- 18 Satoru Fujishige. *Submodular Functions and Optimization*. Elsevier, 2nd edition, 2005.
- 19 Dion Gijswijt and Guus Regts. On the Caratheodory rank of polymatroid bases, 2010. arXiv:1003.1079.
- 20 Ashish Goel and Adam Meyerson. Simultaneous Optimization via Approximate Majorization for Concave Profits or Convex Costs. *Algorithmica*, 44(4):301–323, 2006. doi:10.1007/s00453-005-1177-7.
- 21 Ashish Goel, Adam Meyerson, and Serge A. Plotkin. Approximate majorization and fair online load balancing. *ACM Transactions on Algorithms*, 1(2):338–349, 2005. doi:10.1145/1103963.1103970.
- 22 Godfrey Harold Hardy, John Edensor Littlewood, and George Pólya. *Inequalities*. Cambridge university press, 1934.
- 23 Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.
- 24 Jovan Karamata. Sur une inégalité relative aux fonctions convexes. *Publications de l’Institut Mathématique*, 1(1):145–147, 1932.

- 25 Jon M. Kleinberg and Manish Raghavan. Selection problems in the presence of implicit bias. In *Proceedings of the 9th Conference on Innovations in Theoretical Computer Science Conference (ITCS)*, pages 33:1–33:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ITCS.2018.33.
- 26 Amit Kumar and Jon M. Kleinberg. Fairness measures for resource allocation. *SIAM Journal on Computing*, 36(3):657–680, 2006. doi:10.1137/S0097539703434966.
- 27 Nimrod Megiddo. Optimal flows in networks with multiple sources and sinks. *Mathematical Programming*, 7(1):97–107, December 1974. doi:10.1007/BF01585506.
- 28 Alexander Schrijver. *Combinatorial Optimization: Polyhedra and Efficiency*, volume B. Springer, January 2003.
- 29 Zeyu Shen, Zhiyi Wang, Xingyu Zhu, Brandon Fain, and Kamesh Munagala. Fairness in the assignment problem with uncertain priorities. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 188–196. ACM, 2023. doi:10.5555/3545946.3598636.
- 30 Ashudeep Singh, David Kempe, and Thorsten Joachims. Fairness in ranking under uncertainty. In *Proceedings of the 35th Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 11896–11908, 2021. URL: <https://proceedings.neurips.cc/paper/2021/hash/63c3ddcc7b23daa1e42dc41f9a44a873-Abstract.html>.
- 31 Arie Tamir. Least Majorized Elements and Generalized Polymatroids. *Mathematics of Operations Research*, 20(3):583–589, 1995. doi:10.1287/moor.20.3.583.
- 32 Wei Tang, Haifeng Xu, Ruimin Zhang, and Derek Zhu. Intrinsic robustness of prophet inequality to strategic reward signaling, 2024. doi:10.48550/arXiv.2409.18269.
- 33 Arthur F. Veinott. Least d -Majorized Network Flows with Inventory and Statistical Applications. *Management Science*, 17(9):547–567, 1971. doi:10.1287/mnsc.17.9.547.
- 34 Haifeng Xu, Zinovi Rabinovich, Shaddin Dughmi, and Milind Tambe. Exploring information asymmetry in two-stage security games. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pages 1057–1063. AAAI Press, 2015. doi:10.1609/AAAI.V29I1.9290.

A Proof of Lemma 11

Proof. We define the saturation function $g : 2^E \rightarrow \mathbb{R}_+$ associated with $\mathcal{U}$ as:

$$g(S) = \max_{u \in \mathcal{U}} \sum_{i \in S} u_i = \max_{x \in \mathcal{X}^*} \sum_{i \in S} v_i x_i,$$

where v_i is the i^{th} coordinate of $\vec{v}$ and $u_i = v_i x_i$ is the i^{th} coordinate of u . This function is trivially monotone. We will show that g is submodular. We do so by deriving a closed-form expression for $g(S)$. Subsequently, we will show that $\mathcal{U} = \mathcal{B}(g)$, completing the proof.

By Lemma 6, the vertices of the optimal face $\mathcal{X}^*$ are generated by the polymatroid greedy algorithm for the objective vector $\vec{v}$, with different outcomes arising from different tie-breaking orders for agents with the same value v_i . The value $g(S)$ is therefore achieved at the vertex obtained by running the greedy algorithm with a tie-breaking rule that prioritizes maximizing the utility from the set S .

Let the agents E be partitioned into blocks $E_1, E_2, \dots, E_k$ where all agents in a block E_j have the same value v_j , and $v_1 > v_2 > \dots > v_k > 0$. The greedy algorithm proceeds through these blocks sequentially. To find the value of $g(S)$, we define a specific permutation π_S as follows: For each block E_j , agents in $S_j = S \cap E_j$ are processed first. Agents in $E_j \setminus S_j$ are processed next. Within these subsets, any fixed arbitrary order is used. The allocation vector for this permutation is $x(\pi_S)$, and $g(S) = \sum_{i \in S} v_i x_i(\pi_S)$. The total utility from the agents in S is therefore:

$$g(S) = \sum_{j=1}^k \sum_{i \in S_j} v_j x_i(\pi_S) = \sum_{j=1}^k v_j \left(\sum_{i \in S_j} x_i(\pi_S) \right).$$

We now analyze the inner sum for a single block j . Let $P_{<j} = E_1 \cup \dots \cup E_{j-1}$ be the set of all agents in higher-value blocks. Let the processing order for agents in S_j be $s_1, s_2, \dots, s_m$. Using the greedy allocation for the priority rule discussed above, we have:

$$\begin{aligned} \sum_{i \in S_j} x_i(\pi_S) &= \sum_{j=1}^m x(s_j) = \sum_{j=1}^m (f(P_{<j} \cup \{s_1, \dots, s_j\}) - f(P_{<j} \cup \{s_1, \dots, s_{j-1}\})) \\ &= f(P_{<j} \cup S_j) - f(P_{<j}). \end{aligned}$$

Substituting this back into the expression for $g(S)$, we arrive at the closed-form formula:

$$g(S) = \sum_{j=1}^k v_j [f(P_{<j} \cup (S \cap E_j)) - f(P_{<j})].$$

Let $g_j(S) = v_j [f(P_{<j} \cup (S \cap E_j)) - f(P_{<j})]$. To show this function is submodular, we only need to note that $h_j(S) = f(P_{<j} \cup (S \cap E_j))$ is submodular. Since each $g_j(S)$ is submodular, their sum $g(S) = \sum_j g_j(S)$ is also submodular.

Let $\mathcal{B}(g)$ now denote the base polytope of the polymatroid with rank function g . We will now show that $\mathcal{U} = \mathcal{B}(g)$, completing the proof.

First, by definition, $\mathcal{X}^*$ is the set of allocations $x \in \mathcal{P}$ that maximize $v \cdot x$. Let this maximum welfare be $W_{\max}$. Thus, for any $x \in \mathcal{X}^*$, we have $v \cdot x = W_{\max}$. Now, consider any $u \in \mathcal{U}$. By definition, $u_i = v_i x_i$ for some $x \in \mathcal{X}^*$, so that $\mathcal{U}$ lies on the hyperplane $\{z \in \mathbb{R}^n \mid \sum z_i = W_{\max}\}$. Further, for any $u \in \mathcal{U}$ and any $S \subseteq E$, we have $\sum_{i \in S} u_i \leq g(S)$ by definition of $g(S)$. These two observations imply $\mathcal{U} \subseteq \mathcal{B}(g)$. Further, $\mathcal{U}$ is convex, since $\mathcal{X}^*$ is a face of $\mathcal{P}$ and is hence convex.

We will finally show that $\mathcal{B}(g) \subseteq \mathcal{U}$ by showing that any vertex of $\mathcal{B}(g)$ corresponds to a feasible realization of utilities; since $\mathcal{U}$ is convex, this will imply any interior point of $\mathcal{B}(g)$ also lies in $\mathcal{U}$. By Lemma 7, any vertex of $\mathcal{B}(g)$ can be obtained by ordering the elements of E ; relabel them as $1, 2, \dots, n$ in this ordering, and setting $u_i = g([i]) - g([i-1])$. Let $S = [i-1]$, and let $i \in E_j$ as defined above. When we add i to S , it can be checked that $u_i = v_j \cdot (f(P_{<j} \cup \{i\} \cup (S \cap E_j)) - f(P_{<j} \cup (S \cap E_j)))$. This corresponds to the receiver assigning $x_i = f(P_{<j} \cup \{i\} \cup (S \cap E_j)) - f(P_{<j} \cup (S \cap E_j))$, that is, placing i next in the tie-break ordering for E_j after the elements of $P_{<j}$ and $S \cap E_j$, and allocating greedily. Therefore, any vertex of $\mathcal{B}(g)$ corresponds to a feasible realization of utilities by some tie-breaking rule of the receiver. This implies $\mathcal{B}(g) \subseteq \mathcal{U}$. Therefore, $\mathcal{U}$ is the base polytope $\mathcal{B}(g)$ for the submodular function g . $\blacktriangleleft$

B Example Illustrating Signaling and Fairness

We present an example to demonstrate the simultaneous failure of naive information policies to achieve a good approximation ratio, and the power of a carefully designed signaling policy. We consider $n+1$ agents $\{0, 1, \dots, n\}$, with the constraint that at most one agent can be selected. The polymatroid is therefore the set of allocation vectors (probability of selection) that have ℓ_1 norm at most one. The receiver selects the agent with highest posterior mean, using a randomized tie-breaking rule specified by the sender.

Let $q = 1/\sqrt{n}$. Agent 0 has a deterministic quality $v_0 = 2 - q$, while agents $i \in \{1, \dots, n\}$ have i.i.d. quality that is $1/q$ with probability q and 1 otherwise, so all agents have the same prior mean of $2 - q$. The two baseline policies illustrate the following trade-off:

- In the no-revelation policy, we assume the receiver allocates to an agent uniformly at random, since their posterior means are identical. This results in a total welfare of $2 - q = O(1)$, while the max-min fair value achieved is $\frac{2-q}{n+1} = \Theta(1/n)$.
- In the full-revelation policy, agent 0 is chosen only when all other agents have value 1, while happens with probability $O(e^{-\sqrt{n}})$. The social welfare is now $\Theta(\sqrt{n})$, which is a factor of $\Theta(\sqrt{n})$ larger than that of no-revelation. On the other hand, its max-min utility is now $O(e^{-\sqrt{n}})$, a super-polynomial factor worse than that of no-revelation.

Note that in our example, $V = 1/q$, so that the approximation ratio for majorization in Theorem 1 is $O(\log n/\epsilon)$. Clearly, the above two policies do not achieve this.

We now construct a policy that is simultaneously a constant-factor approximation to the social welfare of full-revelation and the max-min fair value of no-revelation. The sender designs a scheme for each agent $i \in \{1, \dots, n\}$: if its true value is $1/q$, send a “HIGH” signal with a small probability $p = 1/(nq)$; otherwise, send a “LOW” signal. This ensures the probability of any single agent sending a HIGH signal is exactly $q \cdot p = 1/n$. The receiver’s posterior means are then:

- $\mathbb{E}[v_i | \text{HIGH}_i] = 1/q = \sqrt{n}$, since the HIGH signal is only ever sent in the high-value state.
- For the LOW signal, we use Bayes’ rule:

$$\begin{aligned} \mathbb{E}[v_i | \text{LOW}_i] &= \frac{\Pr(\text{LOW} | v_i = \frac{1}{q}) \Pr(v_i = \frac{1}{q}) \cdot \frac{1}{q} + \Pr(\text{LOW} | v_i = 1) \Pr(v_i = 1) \cdot 1}{\Pr(\text{LOW})} \\ &= \frac{(1-p)q \cdot \frac{1}{q} + 1 \cdot (1-q)}{(1-p)q + (1-q)} = \frac{1-p+1-q}{1-pq} = \frac{2-q-p}{1-pq}. \end{aligned}$$

Substituting $p = 1/(nq)$, for large n this posterior mean is $\frac{2-1/\sqrt{n}-1/n^{1.5}}{1-1/n}$, which is slightly smaller than $2 - q$.

The receiver’s strategy is as follows: if any HIGH signals are received, select one of these agents (posterior mean $= \sqrt{n}$); if all signals are LOW (an event with constant probability for large n), select agent 0 (value $2 - q$) over the others (posterior ≈ 2). This policy achieves an expected social welfare of $\Theta(\sqrt{n})$, which is within a constant factor of the optimal welfare. At the same time, it guarantees a max-min utility of $\Theta(1/n)$, as agents $1, 2, \dots, n$ are selected with probability $\Theta(1/n)$ each. This single policy is therefore a constant-factor approximation to both the optimal social welfare (achieved by full revelation) and the optimal max-min utility (which is $\Theta(1/n)$).

C Impossibility of Majorization with Non-Polyamatroidal Constraints

We construct a simple, deterministic allocation problem to show that for certain non-polyamatroidal constraints, the approximation factor for majorization must grow at least linearly with the number of agents, showing that Theorem 1 cannot be generalized to arbitrary constraint sets, and requires special properties of polymatroids. Our counterexample holds when $V = 1$ and all value distributions D_i are deterministic, so that no signaling is required.

Consider n agents, each with a deterministic value of $v_i = 1$. Utility is therefore equal to allocation. The set of feasible allocations $\mathcal{P}$ is the convex hull of n vectors $\{u^{(1)}, \dots, u^{(n)}\} \subset \mathbb{R}^n$. For a large constant $M > n$, the vector $u^{(j)}$ is defined by its components $u_k^{(j)} = M^j$ if $k \geq j$ and $u_k^{(j)} = 0$ if $k < j$. Any feasible allocation is a convex combination $x = \sum_{j=1}^n p_j u^{(j)}$ for some probability vector p . A crucial property of this construction is that any feasible allocation x is sorted: $x_k = \sum_{j=1}^k p_j M^j = x_{k-1} + p_k M^k \geq x_{k-1}$. Thus, the j smallest utilities are simply the first j components of the allocation vector, and $Q_j(x) = \sum_{k=1}^j x_k$.

First, we find the optimal policy for each prefix sum objective. The objective $Q_j(x)$ is a linear function of the probabilities p , so its maximum must be achieved at a vertex of the probability simplex, i.e., by a pure policy $p_k = 1$ for some k . If we choose the policy $p_k = 1$, the allocation is $x = u^{(k)}$, and the prefix sum is $Q_j(u^{(k)}) = (j - k + 1)M^k$ if $j \geq k$, and 0 otherwise. Since $M > n$, this value is maximized over $k \in \{1, \dots, j\}$ when $k = j$. Thus, the optimal policy for maximizing Q_j is the pure strategy $p_j = 1$, and the optimal value is $Q_j^* = M^j$.

Now, let us assume a single policy $x = \sum p_j u^{(j)}$ is β -majorized for some $\beta = o(n)$. This requires $Q_j(x) \geq Q_j^*/\beta = M^j/\beta$ for all $j \in \{1, \dots, n\}$. The exact value of the prefix sum is $Q_j(x) = \sum_{i=1}^j p_i M^i (j - i + 1)$. We can bound this by isolating the dominant term: $Q_j(x) = p_j M^j + \sum_{i=1}^{j-1} p_i M^i (j - i + 1)$. The summation is clearly bounded above by nM^{j-1} (since $\sum p_i \leq 1$ and $j - i + 1 \leq n$). The majorization condition thus implies $p_j M^j + nM^{j-1} \geq M^j/\beta$. Dividing by M^j , we get $p_j + n/M \geq 1/\beta$, which gives the necessary condition $p_j \geq 1/\beta - n/M$ for each $j = 1, \dots, n$:

$$1 = \sum_{j=1}^n p_j \geq \sum_{j=1}^n \left(\frac{1}{\beta} - \frac{n}{M} \right) = \frac{n}{\beta} - \frac{n^2}{M}.$$

Rearranging this gives a lower bound on the required approximation factor: $\beta \geq \frac{n}{1+n^2/M}$. By choosing M to be a sufficiently large polynomial in n (e.g., $M = n^3$), this implies $\beta \geq \Omega(n)$. This contradicts the assumption that β is sub-linear. Therefore, no such policy can exist.

D Discussion on Lemma 11

We now present two pieces of evidence to show the non-triviality of Lemma 11, in that it needs delicate arguments that are tailored to the setting we consider. First, we show that for non-polymatroidal allocation constraints, 1-majorization in the allocation space will not imply majorization in the utility space, to any sub-linear approximation. This shows that our proof crucially requires the polymatroidal structure of the allocation space. We next show that part of our argument is not a generic result for polymatroids, and is specific to the utility polytope we define. In particular, we show that the result $\mathcal{B}(g) \subseteq \mathcal{U}$ is not merely a consequence of the submodularity of g (where g is the saturation function of $\mathcal{U}$), and the result is only true for the specific $\mathcal{U}$ (utility vectors of the receiver-optimal allocation) that we define.

D.1 Impossibility with Majorized Allocation Set

We first show a non-polymatroidal allocation set that has a 1-majorized point, but even with deterministic values of the agents, the corresponding set of utility vectors does not have a point with sub-linear approximation to majorization. This rules out generalizing Lemma 11 to non-polymatroidal constraints.

In the example below, the value distributions D_i are deterministic, so no signaling is required. Consider $2n$ agents, evenly partitioned into n groups $g_1, \dots, g_n$, such that $g_i = \{i, 2n - i\}$. The feasible allocations is the convex hull of the indicator vectors $\{\mathbb{1}[g_1], \dots, \mathbb{1}[g_n]\} \subset \mathbb{R}^n$. The allocation that selects each g_i with probability $1/n$ is a 1-majorized policy.

Let $N = \omega(n^{2n})$. For $1 \leq i \leq n$, let agent i take value $v_i = n^{2i}$ and agent $2n - i$ take value $v_{2n-i} = N - n^{2i}$. Since the total value of any group is N , the set of receiver-optimal allocations is the probability simplex $\{\vec{p} \mid \sum_{i=1}^n p_i = 1\}$, where p_i is the probability of selecting group g_i .

We first derive a lower bound for the sum of the smallest i utilities. Consider the following class of allocations $\{A_i\}_{i=1}^n$. Define A_i to be the allocation that selects g_j with probability $1/n^2$ for $j \neq i$ and g_i with probability $1 - (n-1)/n^2$. Then, the i^{th} smallest utility for A_i is

$$\left(1 - \frac{n-1}{n^2}\right) \cdot n^{2i} \geq n^{2i} - n^{2i-1}.$$

This then lower bounds the sum of the i smallest utilities.

Suppose next that there is an α -majorized policy A for the set of utility vectors, which selects g_i with probability p_i . Then, the sum of i smallest utilities is at most

$$\sum_{j=1}^{i-1} n^{2j} + p_i \cdot n^{2i} \leq \left(p_i + \frac{2}{n^2}\right) \cdot n^{2i}.$$

For A to be α -majorized,

$$\alpha \cdot \left(p_i + \frac{2}{n^2}\right) \cdot n^{2i} \geq n^{2i} - n^{2i-1}.$$

Therefore,

$$\alpha \cdot \left(p_i + \frac{2}{n^2}\right) \geq 1 - o(1), \quad \text{and} \quad p_i \geq \frac{1 - o(1)}{\alpha} - \frac{2}{n^2}.$$

Since this holds for all i and $\sum_{i=1}^n p_i = 1$, we have

$$\frac{n(1 - o(1))}{\alpha} - \frac{2}{n} \leq 1.$$

This implies

$$\alpha \geq \frac{n(1 - o(1))}{1 + \frac{2}{n}} = n(1 - o(1)).$$

Therefore, the best majorization factor for the set of utility vectors grows linearly with the number of agents, despite the existence of a 1-majorized point for the set of allocation vectors.

D.2 Submodular Saturation Functions do not Suffice

We next show that the result $\mathcal{B}(g) \subseteq \mathcal{U}$ is not a general result for arbitrary convex polytopes $\mathcal{U}$ that are constant sum and whose saturation function g is submodular. Here, constant-sum means the sum of coordinates is a constant. Also recall that the saturation function $g(S)$ for a polytope is the maximum over the polytope of the sum of the coordinates in S . We show an example where a constant sum convex polytope $\mathcal{U}$ has submodular saturation function g , but is strictly contained in the base polytope $\mathcal{B}(g)$. This shows the proof of Lemma 11 is delicate in requiring specific properties of the $\mathcal{U}$ that we define.

We start by defining the submodular function g . Let h be defined as $h(0) = 0$, $h(1) = 1$, $h(2) = 1.9$, and $h(3) = 2.5$. This is clearly concave. Let $g(S) = h(|S|)$ for sets S of size at most three. Consider the 3-dimensional base polytope $\mathcal{B}(g)$. This is defined by the constraints:

$$\{(x, y, z) \geq 0 \mid \max\{x, y, z\} \leq 1; \max\{x+y, y+z, x+z\} \leq 1.9; x+y+z = 2.5\}.$$

Projecting onto the (x, y) plane, we obtain the following hexagon:

$$\{(x, y) \geq 0 \mid \max\{x, y\} \leq 1; \min\{x, y\} \geq 0.6; 1.5 \leq x+y \leq 1.9\}.$$

Note that there is a one-to-one mapping between the points in the planar hexagon and the points in $\mathcal{B}(g)$. Further, every edge in the planar hexagon corresponds to one saturation function of $\mathcal{B}(g)$. For instance, the edge $x + y = 1.5$ corresponds to $z = 1$ (corresponding to the set $S = \{3\}$, the dimension for z). Similarly, $x = 0.6$ corresponds to $y + z = 1.5$ (corresponding to the set $S = \{2, 3\}$, the dimensions for y, z). The vertices of the hexagon are of the form (a, b) , where $a \neq b$, and $a, b \in \{1, 0.9, 0.6\}$.

Now consider some vertex of the hexagon, say $(1, 0.6)$ and the corresponding vertex $v \in \mathcal{B}(g)$. Define $\mathcal{U} \subset \mathcal{B}(g)$ by intersecting the planar hexagon with a halfspace that removes v from $\mathcal{B}(g)$, but preserves the other edges and vertices. For instance, add the halfspace $x - y \leq 0.39$. Note that the resulting projection of the convex polytope $\mathcal{U}$ on the plane is defined by all of the original halfspaces of the planar hexagon, plus the new halfspace. This means that for any subset S of dimensions, the function $g'(S)$ that maximizes the sum of the coordinates in S over $\mathcal{U}$ coincides with the same function for $\mathcal{B}(g)$, which is $g(S)$. But $\mathcal{U} \subset \mathcal{B}(g)$ by construction. This also means $\mathcal{U}$ is constant sum, since $x + y + z = 2.5$ by construction. Therefore, we have an example $\mathcal{U}$ whose saturation function g is submodular, but which lies strictly within $\mathcal{B}(g)$.