Eligibility Mechanisms: Auctions Meet Information Retrieval

Gagan Goel
Google
USA
gagangoel@google.com

Renato Paes Leme
Google
USA
renatoppl@google.com

Jon Schneider
Google
USA
jschnei@google.com

David Thompson
Google
USA
davidrthompson@google.com

Hanrui Zhang
Carnegie Mellon University
USA
hanruiz1@cs.cmu.edu

ABSTRACT

The design of internet advertisement systems is both an auction design problem and an information retrieval (IR) problem. As an auction, the designer needs to take the participants incentives into account. As an information retrieval problem, it needs to identify the ad that it is the most relevant to a user out of an enormous set of ad candidates. Those aspects are combined by first having an IR system narrow down the initial set of ad candidates to a manageable size followed by an auction that ranks and prices those candidates.

If the IR system uses information about bids, agents could in principle manipulate the system by manipulating the IR stage even when the subsequent auction is truthful. In this paper we investigate the design of truthful IR mechanisms, which we term eligibility mechanisms. We model it as a truthful version of the stochastic probing problem. We show that there is a constant gap between the truthful and non-truthful versions of the stochastic probing problem and exhibit a constant approximation algorithm. En route, we also characterize the set of eligibility mechanisms, which provides necessary and sufficient conditions for an IR system to be truthful.

CCS CONCEPTS

• **Theory of computation → Algorithmic game theory and mechanism design**.

KEYWORDS

mechanism design, online advertising

ACM Reference Format:
Gagan Goel, Renato Paes Leme, Jon Schneider, David Thompson, and Hanrui Zhang. 2023. Eligibility Mechanisms: Auctions Meet Information Retrieval. In *Proceedings of The Web Conference (WWW'23)*. ACM, New York, NY, USA, 12 pages. https://doi.org/XXXXXXX.XXXXXXX

WWW'23, April 30 – May 4, 2023, Austin, TX

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

The design of internet advertisement systems is both an *auction design* problem and an *information retrieval* (IR) problem. As an auction, the designer needs to take the participants' incentives into account. As an information retrieval problem, it needs to identify the ad that is the most relevant to a user out of an enormous set of ad candidates. The paradigmatic problem in information retrieval is organic web search. In this problem, the corpus of possible results is enormous and needs to be narrowed down to a small set by the application of successive filtering algorithms, until we have a small enough set to rank and display to users. An ad system is not very different. We have again an enormous corpus of ad candidates that we need to filter to a small set that then participates in the auction. Filters applied early in the funnel must be very fast as they are applied to a huge number of candidates. Filters applied later on can be more computationally costly. To give some examples, one may early on exclude all ads whose language is different from the user's language, and later on exclude an ad because a machine learning model tells us that the probability of a user clicking on this ad is very low.

While there is a large body of work on both the auction aspect [10, 28] and the information retrieval aspect [4, 5, 9] of ads, very little is known about their interaction. In fact, we can decouple both problems by preventing the IR system from using any bid information so that advertisers can't influence the filtering directly. The auction, then, ranks and prices the set of candidates returned by IR, only then using the bid information.

In this paper we investigate how to design an incentive compatible information retrieval system that can benefit by taking bid information into account. To make this problem concrete, consider the problem of maximizing welfare in a click-through auction, where n buyers compete for a single ad slot. Each buyer i is described by their bid v_i representing the value for a click, which is reported by the buyer and a click-through rate (CTR) r_i specifying the probability of a click. The ideal solution for the auctioneer is to choose the ad maximizing the product $r_i v_i$ and then price according to a second price auction.

In practice, the CTRs r_i are computed by an expensive machine learning model, by taking into account all available contextual information about the user, the ad creative and the landing page. Because the auction happens in real-time, the auctioneer can't afford to apply this model on all possible ad candidates. Instead, what happens in practice is that the ad system first computes rough estimates $\tilde{r}_i$ of the CTRs based on a few salient features. Given these estimates, the auctioneer narrows the set of buyers down to a handful of potential winners, by, for example, sorting buyers by $v_i\tilde{r}_i$ and picking the top k buyers. Restricted to these potential winners, the auctioneer then performs a refined procedure to obtain more accurate estimates of the CTRs, which are then used to run a second-price auction among potential winners and decide the final winner. A more detailed discussion of this design in practice can be found in [18] for Facebook[1] and in [29] for Alibaba.

Truthful stochastic probing. The above procedure that determines the set of potential winners can be viewed as a stochastic probing problem: each buyer i is associated with a value v_i and a CTR *distribution* R_i. The actual CTR of each ad i is given by $r_i \sim R_i$, independent across buyers. The distribution R_i encodes the rough estimate given by a weaker machine learning model while the sample r_i encodes the refined estimate.

Knowing each v_i and R_i, the auctioneer can probe a subset $K \subseteq [n]$ of $|K| = k < n$ buyers. After probing, the auctioneer learns the value of $r_i \sim R_i$ for all $i \in K$ and then can select the buyer $i \in K$ with the highest $v_i r_i$. The expected welfare is then $\mathbb{E}_{r_i \sim R_i}[\max_{i\in K} v_i r_i]$, which the auctioneer seeks to maximize by choosing K appropriately.

If we ignore incentives, this is precisely the ProbeMax problem ([8], [11], [22] and [25]). While the problem is known to be NP-hard, Segev and Singla [25] have recently shown that it admits an efficient polynomial-time approximation scheme (EPTAS). Those algorithms, however, are not incentive compatible, in the sense that agents may misreport their private value to manipulate the probing procedure to exclude competitors from the later auction. To see why this is the case, consider the following situation: suppose the seller wants to implement the "first-best" allocation, which, given v, first probes a set K that maximizes $\mathbb{E}[\max_{i\in K} v_i r_i]$, and then upon the realization of the CTRs, chooses the buyer in K with the highest CTR as the winner. If a buyer is better informed about the other buyers' CTRs than the seller before the probing stage, then this buyer might be motivated to misreport his value to exclude the buyers with high CTR from the probed set K. In Theorem 1 we show a concrete instance of such an example and argue that the seller cannot prevent such manipulation by charging appropriate payments based on the *seller's predictions* (rather than the realized values) of the CTRs.

The above argument illustrates a gap between the first-best welfare, i.e., the optimal welfare that can be guaranteed when all values are known to the auctioneer, and the second-best welfare, i.e, the maximum welfare that can be guaranteed when the values need to be reported by strategic agents. This brings us to the main goal of this paper, which is to *formally understand this gap and the computational aspects thereof.*

1.1 Our Results

First-best is not achievable. We formulate a combined problem where a seller starts with a very large collection of ad candidates with rough information about their quality (CTRs) and needs to select a subset of the agents to probe in order to learn their quality. When defining the notion of truthfulness for this setting we apply the philosophy that inaccuracies in ML predictions may affect the performance of the mechanism but should not harm incentives. This has been a guiding principle behind the design of auctions in online advertising. In our setting, this leads to the notion of ex-post truthfulness: the combined auction should be truthful regardless of what information or belief the buyers have about the CTRs. Our first result is a lower bound showing a constant gap between the truthful and non-truthful version of the problem.

Notion of eligibility mechanisms and their characterization. Our problem can be thought of two separate stages: the eligibility stage where we select/probe the ads that will participate in the auction and the auction stage where we rank and price candidates. Motivated by the modular nature of mechanism design in practice, we introduce a framework to decouple the design of the probing stage and the subsequent auction stage. We say that a probing mechanism is truthful when it *remains truthful when composed with any downstream auction mechanism that is truthful by itself.* We call such mechanisms "eligibility mechanisms". We provide a complete characterization of eligibility mechanisms: such a mechanism (in addition to being monotone) must satisfy the property that for any agent, whenever this agent is chosen by the mechanism, the action of this agent cannot affect which other agents are also chosen. In other words, no agent can choose their competitors in the next stage by misreporting in the current stage, unless they cannot enter the next stage in the first place.

The practical relevance of this characterization is that it enables practitioners to design information retrieval systems that may use bid information without harming incentives. From a theory standpoint, this characterization enables the subsequent results on approximation mechanisms.

Constant-factor eligibility mechanisms. We study the problem where the mechanism can probe k agents and then choose $m < k$ agents to allocate in the auction. Here, m corresponds to the number of ad slots to be allocated (for simplicity, we assume the ad slots are identical). We then give an efficient eligibility mechanism that approximates the first-best welfare within a factor of $2e/(e-1)$ when $m = 1$; this factor goes to 2 when m goes to infinity, roughly at a rate of $2 + O(1/\sqrt{m})$. This implies that the gap between the first-best welfare and the second-best welfare is constant.

Moreover, we give another efficient eligibility mechanism with an approximation ratio of $O(\log(k/m) + 1)$. While this ratio is generally super-constant, we will later see that the unique properties of this mechanism makes it useful in the more general multi-stage setting. Technically, both our mechanisms are based on an ex-ante relaxation of the probing problem, which has proved useful in posted-price mechanisms [2, 6], prophet inequalities [17, 24] and team formation [20]. The main technical innovation is the use of this relaxation to design a truthful mechanism: we approximate the optimal solution to the ex-ante relaxation by another solution to the same relaxation, which can be computed in a truthful way.

Multi-stage eligibility mechanisms. We finally consider a more general setting where the mechanism gradually obtains more and more refined information about fewer and fewer agents by probing

[1][18] say: *"In order tackle a very large number of candidate ads per request, where a request for ads is triggered whenever a user visits Facebook, we would first build a cascade of classifiers of increasing computational cost"*

in multiple stages. This model is inspired by the successive levels of filtering that ad candidates go through in practice (see footnote 1, for example). In this model, each agent i is associated with a Markov chain $s_{t,i}$ for $t = 1, 2, \dots, T+1$ and the CTR is a function of the last state $s_{T+1,i}$. Given $k_1 \ge k_2 \ge \dots \ge k_T$, the mechanism can probe the agents in T stages. In the t-th stage, the mechanism can probe k_t agents out of the k_{t-1} agents who are still active (any agent who is not probed immediately becomes inactive). After agent i is probed in stage t, the mechanism learns $s_{t,i}$. The goal is again to maximize the (expected) welfare among the agents who are still active after all T stages. To showcase the modularity of our approach, we analyze the mechanism obtained by stacking T copies of our $O(\log(k/m)+1)$-approximate single-stage eligibility together.

1.2 Further Related Work

Our results are along the line of research on optimization under uncertainty. In particular, the stochastic probing problems that we consider (and generalizations thereof) have been studied in environments without strategic behavior [1, 3, 12–16, 27]. For a comprehensive overview, see the dissertation by Singla [26] and references therein. The main difference between our results and prior work on stochastic probing is that we consider probing mechanisms that are truthful, while prior work does not take strategic behavior into consideration.

Closest to our model is the paper by [29] on two-stage auction design, who also study the problem of selecting a subset of ad candidates to enter the auction. The focus of their paper is on machine learning techniques for the first stage. While they briefly discuss incentives, their incentive analysis is in a non-standard utility model where agents are indifferent to payments as long as they are below their value. Our paper's main focus is on incentives in the standard quasi-linear model.

2 THE COMBINED PROBING+AUCTION PROBLEM AND A LOWER BOUND

In this section, we formally define the combined problem of designing a first-stage probing mechanism and a second-stage auction mechanism in order to maximize the social welfare resulted from the overall allocation. We consider mechanisms that are "ex-post truthful", meaning that reporting truthfully is a dominant strategy *fixing the CTRs* r_i. As we argue below, this notion of truthfulness captures robustness against potentially inaccurate ML estimates of the CTRs, as well as misaligned beliefs between the seller and the buyers. We then give a lower bound which illustrates a gap between the first-best welfare and the best welfare that can be achieved by an ex-post truthful combined mechanism. This motivates our investigation on eligibility mechanisms, as discussed in later sections.

Setting. Throughout this paper, we consider single-parameter settings with n agents[2] $[n]$, where the private information of each agent i is a single value $v_i \in \mathbb{R}_+$. Each agent has also a click-through rate r_i that is drawn from a publicly known distribution R_i over $[0,1]$. We use $R = (R_1, \dots, R_n)$ to denote the product distribution of each R_i. The mechanism knows R_i but in order to observe r_i, it needs to probe[3] agent i. We consider *nonadaptive* probing, which means the mechanism must choose a subset $K \subseteq [n]$ of agents beforehand, and then probe all agents in K simultaneously.[4] After the agents are probed, the mechanism decides the final winner(s) based on $v^K = \{v_i\}_{i \in K}$ and $r^K = \{r_i\}_{i \in K}$, and charges payments for agents in K.

Formally, a combined mechanism consists of a probing mechanism f and a winner determination mechanism $(\{x^K, p^K\}_K)$. Below we define the two parts respectively.

For the probing mechanism, let $\mathcal{F} \subseteq 2^{[n]}$ be a feasibility set representing a set of agents that can be probed simultaneously. For most of the paper we focus on sets of the type $\mathcal{F} = \{K \subseteq [n] \mid |K| \le k\}$, but some of our results are able to accommodate more sophisticated settings such as arbitrary matroid constraints. Fixing the CTR distributions R_i associated with the agents (which are given as prior knowledge), a probing mechanism is a mapping $f : \mathbb{R}_+^n \to \Delta(\mathcal{F})$ associating a vector of bids $(v_1, \dots, v_n)$ to a distribution over feasible subsets $K \in \mathcal{F}$ to probe given those bids. In the rest of the paper, unless otherwise specified, we focus on truthful mechanisms where the bids are always the same as the true values.

The winner determination mechanism is parametrized by the set K probed by the probing mechanism. Fixing K, $x^K : \mathbb{R}_+^K \times [0,1]^K \to [0,1]^K$ maps the vector of bids v^K and the vector of CTRs r^K (which becomes available after probing) of agents in K to an allocation vector $x^K(v^K, r^K)$, where for each $i \in K$, $x_i^K(v^K, r^K)$ is the fraction of the item agent i receives. Similarly, $p^K : \mathbb{R}_+^K \times [0,1]^K \to \mathbb{R}_+^K$ maps v^K and r^K to the payment vector, where for each $i \in K$, $p_i^K(v^K, r^K)$ is the amount agent i pays. Note that we allow the winner determination mechanism to behave differently on different subsets of agents. That is, the winner determination mechanism is actually given by a collection of allocation and payment rules $\{(x^K, p^K)\}_K$, one for each feasible subset K.

Given a probing mechanism f and a winner determination mechanism $\{(x^K, p^K)\}_K$, the allocation rule of the combined mechanism $x^f : \mathbb{R}_+^n \times [0,1]^n \to [0,1]^n$ is defined such that for any $v \in \mathbb{R}_+^n$, $r \in [0,1]^n$, and $i \in [n]$,

$$x_i^f(v, r) = \mathbb{E}_{K \sim f(v)}[x_i^K(v^K, r^K)] \tag{1}$$

where $x_i^K \equiv 0$ whenever $i \notin K$. The payment rule of the combined mechanism p^f is defined similarly. In particular, it is possible that (x^K, p^K) for some K is not truthful on its own, but the combined mechanism is still truthful.

Ex-post truthfulness. We focus on combined mechanisms where truthful reporting is a dominant strategy for every realization of the CTRs r. Formally, a combined mechanism (x^f, p^f) is ex-post truthful[5] , if for any $i \in [n]$, $v \in \mathbb{R}_+^n$, $r \in [0,1]^n$ and $v_i' \in \mathbb{R}_+$,

$$x_i^f(v, r) \cdot v_i - p_i^f(v, r) \ge x_i^f(v_i', v_{-i}, r) \cdot v_i - p_i^f(v_i', v_{-i}, r).$$

[2]We use the terms agent, buyer and ad candidate interchangeably.

[3]In practice, probing corresponds to running an expensive ML model to predict the agent's click-through rate.

[4]It is known that probing adaptively can only increase the welfare by a constant factor [26].

[5]The term "ex-post" often means "after *types* are revealed", which, notably, is different from its meaning here: "after *CTRs* are revealed".

In light of Myerson's characterization [23], we focus on the parametrized allocation rule $\{x^K\}_K$, since there exists a parametrized payment rule $\{p^K\}_K$ such that the combined mechanism is ex-post truthful, if and only if the combined allocation rule is non-decreasing. That is, x^f can be implemented truthfully, if and only if for each $i \in [n]$, $v_{-i} \in \mathbb{R}_+^{[n]\setminus\{i\}}$, and $r \in [0,1]^n$, $x^f(v_i, v_{-i}, r)$ is non-decreasing in v_i.

Ex-post truthfulness and robustness. Robust auctions with ML predictions must gracefully handle inaccurate predictions as well as misalignment between agents' beliefs and the prediction. For example, the CTR prediction may be inaccurate in a slice of the inventory where data is sparser and for some segments, the agents themselves may have different beliefs about what the CTR is.

For that reason it is important to design auctions that remain truthful even if the predictions are inaccurate. Note that a pay-per-click second price auction that chooses the winner using the product of CTR and bid is truthful no matter how off the prediction is. A bad ML model can hurt efficiency but not incentives.

For this practical reason we choose ex-post truthfulness instead of just truthfulness in expectation over the CTRs. In fact, it is not hard to see that a combined mechanism remains truthful regardless of the agents' prior beliefs, if and only if it is ex-post truthful. This is because if the former is true, then the combined mechanism must be truthful when all agents share any single-point prior belief, which means the combined mechanism is ex-post truthful. On the other hand, if the mechanism is ex-post truthful, then conditioned on the realized CTRs, it is still truthful. Then taking the expectation over each agent's prior belief, the mechanism must also be truthful given any prior beliefs of the agents. This is especially desirable in the context of ad auctions, because the ML algorithms used to predict the CTRs are almost never perfectly accurate, neither are the agents' predictions of the CTRs. In other words, there is almost certainly a misalignment between the mechanism's prior and the agents'. In such cases, ex-post truthfulness ensures that the mechanism is still truthful, even under arbitrary misalignment of prior beliefs.

A lower bound. Ex-post truthfulness is desirable. However, as we show below, it is impossible to achieve the first-best welfare using ex-post truthful combined mechanisms.

Theorem 1. *In a setting with $n = 3$ agents, out of which 2 are probed, i.e., $\mathcal{F} = \{K \subseteq [n] \mid |K| \le 2\}$, there exists CTR distributions such that no ex-post truthful combined mechanism approximates the first-best welfare within a factor of $\frac{23}{22} \approx 1.045$.*

Proof. Let the CTRs be such that $r_1 = 1$,

$$r_2 = \begin{cases} 1, & \text{with probability } 1/3 \\ 0, & \text{with probability } 2/3 \end{cases}, \; r_3 = \begin{cases} 1, & \text{with probability } 1/10 \\ 0, & \text{with probability } 9/10 \end{cases}.$$

Moreover, let $v_2 = 2$ and $v_3 = 3$. Fix any ex-post truthful mechanism x^f formed by f and $\{x^K\}_K$. Consider the behavior of f on $v = (1, v_2, v_3) = (1, 2, 3)$ and $v' = (2, v_2, v_3) = (2, 2, 3)$. First observe that the first-best mechanism (which is not ex-post truthful) probes $\{1, 2\}$ on v, and $\{1, 3\}$ on v'. After probing, the first-best mechanism allocates to agent 1 if and only if the other agent probed has CTR 0. Now consider the behavior of x^f when $r = (1, 0, 1)$. In order for x^f to be ex-post truthful, we must have

$$x_1^f(v, r) \le x_1^f(v', r).$$

Suppose the above is the only incentive constraint that we have (which can only make the problem easier). Then to maximize the objective, we would like to make $x_1^f(v, r)$ as large as possible, and $x_1^f(v', r)$ as small as possible. As a result, without loss of generality we can assume that

$$x_1^f(v, r) = x_1^f(v', r) = t$$

for some $t \in [0, 1]$. Now recall that x^f is the composition of f and $\{x^K\}_K$, and in particular, f cannot depend on the realization of the CTRs, and each x^K cannot depend on the realization of any agent i's CTR if $i \notin K$. As such, the best way that x^f can satisfy the above incentive constraint is to do the following:

- On v, probe $\{1, 2\}$. If $r_2 = 1$, then allocate to 2. Otherwise, allocate to 1 with probability t, and to 2 with probability $1 - t$. This gives welfare $\frac{2}{3} + \frac{2t}{3}$. For comparison, the first-best welfare on v is $\frac{4}{3}$.
- On v', probe $\{1, 3\}$. If $r_3 = 0$, then allocate to 1. Otherwise, allocate to 1 with probability t, and to 3 with probability $1 - t$. This gives welfare $\frac{21}{10} - \frac{t}{10}$. For comparison, the first-best welfare on v' is $\frac{21}{10}$.

To achieve the best approximation ratio possible, we choose the optimal t where

$$\frac{2/3 + 2t/3}{4/3} = \frac{21/10 - t/10}{21/10},$$

which gives $t = \frac{21}{23}$, and the approximation ratio of x^f is $\frac{23}{22}$.

One may wonder if it is possible to do better by probing different sets than what first-best probes. In particular, on v, it might appear beneficial to sometimes probe $\{1, 3\}$ and allocate to 3 with some probability. However, since the best welfare from probing $\{1, 3\}$ on v is $\frac{6}{5}$, one cannot hope for a better ratio than $\frac{4/3}{6/5} = \frac{10}{9} > \frac{23}{22}$ by probing $\{1, 3\}$ on v. Similarly, one also would not want to probe a different set on v' (although that might allow more flexibility on v), because that would make the approximation ratio on v' strictly worse than $\frac{23}{22}$, rendering any potential gain elsewhere meaningless. □

Remark on in-expectation truthful mechanism. We also remark that it is possible to achieve the first-best welfare using an in-expectation truthful mechanism over the random realization of the CTRs. In fact, let $(f, \{x^K\}_K)$ be (the allocation rule of) the first-best mechanism. Then for any agent i, the allocation rule $v \mapsto \mathbb{E}_{r \sim R}[x_i^{f(v)}(v, r)]$ is non-decreasing in v_i. So there exists a payment rule $p_i(v)$ that truthfully implements this allocation rule in expectation. However, the above lower bound shows that this mechanism can't be made ex-post truthful, and therefore will introduce incentive issues when there are misaligned prior beliefs.

3 ELIGIBILITY MECHANISMS

Having established a gap between the first-best welfare and the best welfare achievable by (ex-post) truthful mechanisms, we now focus on the design of truthful approximation mechanisms. Our first step is to present a framework to decouple the design of probing stage and the winner determination stage. Recall that a mechanism has two components:

- a probing mechanism $f : \mathbb{R}_+^n \to \Delta(\mathcal{F})$, and
- a family of winner determination allocation rules $\{x^K\}_{K\in\mathcal{F}}$.

Using equation (1), we can combine those into a single allocation rule x^f. In the previous section we defined truthfulness of the combined mechanism. Here we will define a notion of truthfulness for each component in a way that it is preserved under composition.

For winner determination mechanisms, this notion is rather standard. Recall that in our setting the allocation rule x^K takes as input a vector of valuations v^K and a vector of CTRs r^K.

Definition 1. A winner determination rule $x : \mathbb{R}_+^n \times [0,1]^n \to [0,1]^n$ is truthfully implementable if there exists a payment rule $p : \mathbb{R}_+^n \times [0,1]^n \to \mathbb{R}_+^n$ such that for every $i \in [n]$, $r \in [0,1]^n$, $v \in \mathbb{R}_+^n$ and $v'_i \in \mathbb{R}_+$ it holds that:

$$x_i(v,r)\cdot v_i - p_i(v,r) \geq x_i(v', v_{-i}, r)\cdot v_i - p_i(v', v_{-i}, r).$$

A family of winner determination mechanisms $\{x^K\}_{K\in\mathcal{F}}$ is truthfully implementable iff for each $K \in \mathcal{F}$, the winner determination rule x^K is truthfully implementable.

For probing mechanisms, our definition is somewhat more abstract and less syntactic. However, as we will see momentarily, this abstract definition admits a rather nice syntactic characterization.

Definition 2. A deterministic function $f : \mathbb{R}_+^n \to \mathcal{F}$ is an *eligibility mechanism* if for any truthfully implementable family of winner determination mechanisms $\{x^K\}_{K\in\mathcal{F}}$ its composition x^f is truthfully implementable. A randomized function $f : \mathbb{R}_+^n \to \Delta(\mathcal{F})$ is an eligibility mechanism, if it randomizes over deterministic eligibility mechanisms.

It is easy to check that randomized eligibility mechanisms also preserve truthfulness. We do not consider "truthful-in-expectation" probing mechanisms in this paper for two reasons: (1) eligibility mechanisms according to our definition are already powerful enough, and (2) truthful-in-expectation" mechanisms may not be as robust to strategic behavior. For the rest of the paper we will focus on mechanisms that are compositions of an eligibility mechanism with a truthful winner determination mechanism. This decoupling is very appealing from a practical standpoint: typically auction and information retrieval systems are designed with different goals in mind by teams with different expertise. As long as both components satisfy their respective notions of truthfulness, they can be designed and modified independently.

For this definition to be useful, however, we should provide a clean characterization of eligibility mechanism. This is what we do in the next theorem, which says that a deterministic eligibility is truthful if and only if it is monotone, and no agent chosen by the mechanism can affect which other agents are chosen. By extension, this also characterizes randomized eligibility mechanisms.

Theorem 2. *A deterministic mechanism $f : \mathbb{R}_+^n \to 2^{[n]}$ is an eligibility mechanism, if and only if for any $i \in [n]$, $v \in \mathbb{R}_+^n$, and $v'_i \geq v_i$,*

$$i \in f(v) \implies f(v) = f(v'_i, v_{-i}).$$

The proof of the theorem, as well as all other missing proofs, is deferred to the appendix.

Examples of eligibility mechanisms. Theorem 2 offers a guide for practitioners on how to use bid information when designing information retrieval systems. We now give a few examples of mechanisms satisfying the conditions in the theorem:

- bid-free algorithms: any probing mechanism that doesn't use information about the bids. For example, an algorithm that filters all ad candidates whose language is different from the user's language.
- Top-k mechanisms: a mechanism that sorts agents by a score $s_i(v_i)$ that depends monotonically on v_i and chooses the k agents with the largest scores. For example, the algorithm that probes the agents with largest $v_i \mathbb{E}_{r'_i\sim R_i}[r'_i]$.
- Selection among disjoint groups: given a fixed (bid-independent) partition $T_1, \ldots, T_t$ of $[n]$ and a scoring functions $s_j : \mathbb{R}_+^{T_j} \to \mathbb{R}$, probes the set T_j maximizing $s_j(v^{T_j})$. For example, we can partition agents randomly and then probe the partition with highest $\mathbb{E}_{r'\sim R}[\sum_{i\in T} r'_i v_i]$.

In fact, we show below that a fairly general class of mechanisms, namely maximizers over an arbitrary matroid, are all eligibility mechanisms. Such mechanisms subsume top-k mechanisms as a subclass.

Lemma 1. *Fix an arbitrary matroid feasibility constraint $\mathcal{F}$, and an arbitrary mapping $s_i : \mathbb{R}_+ \to \mathbb{R}_+$ for each agent i. If each s_i is monotonically non-decreasing, then the following mechanism $f : \mathbb{R}_+^n \to \mathcal{F}$ is an eligibility mechanism:*

$$f(v) = \operatorname*{argmax}_{S\in\mathcal{F}} \sum_{i\in S} s_i(v_i).$$

Remarks on payments. Although payments are not the focus of this paper, here we briefly discuss how they can be computed in a modular way when eligibility mechanisms are used for probing. With two or more stages (see Section 5 for the setup), in general, the final payments of the combined mechanism must depend on earlier stages. However, when eligibility mechanisms are used in earlier stages, this dependency is in fact extremely simple: one can show that the overall payment of agent i should simply be $\max_t p_{i,t}$, where $p_{i,t}$ is simply the minimum value required for agent i to be chosen in stage t (conditioned on the randomness of the mechanism). This dependence should not be a problem in decoupling different stages given its sheer simplicity. In particular, no matter what happens in earlier or later stages, each stage i only needs to pass $p_{i,t}$ to the next stage or the module that determines the payments, and modularity can be preserved.

4 CONSTANT-FACTOR ELIGIBILITY MECHANISMS

We complement the impossibility result in Theorem 1 with a constant approximation mechanism. We consider a setting where mechanism can probe any k agents $\mathcal{F} = \{K \subseteq [n] \mid |K| \leq k\}$ and after probing can allocate to at most m agents. The goal of the mechanism is to optimize efficiency of the allocation measured by social welfare, i.e., $\sum_{i\in M} r_i v_i$ for the set M of allocated agents.

We will adopt the framework developed in Section 3 and optimize over eligibility mechanisms $f : 2^{[n]} \to \Delta(\mathcal{F})$ (Definition 2 and Theorem 2). For the winner determination stage, we will select the

m agents with highest $r_i v_i$. Our objective then can be phrased as:

$$\mathrm{obj}(K, v) = \mathbb{E}_{r\sim R}\left[\max_{M\subseteq K:|M|\le m}\sum_{i\in M} v_i r_i\right].$$

We will compare against the first-best benchmark, i.e., the welfare obtained by an algorithm that knows the true values of all agents, but is still constrained in terms of the set of agents it can probe. The approximation ratio α of a given eligibility mechanism f is given by:

$$\alpha = \sup_{v\in\mathbb{R}_+^n}\max_{K\in\mathcal{F}}\frac{\mathrm{obj}(K, v)}{\mathbb{E}_{K'\sim f(v)}[\mathrm{obj}(K', v)]}.$$

Example 1. A natural rule is to select the agents in the probing stage using the best possible estimate of the CTR available before probing, i.e., probe the agents with largest $v_i\mathbb{E}_{r_i\sim R_i}[r_i]$. This mechanism is ex-post truthful by Theorem 2, but its approximation ratio is unbounded. To see that, consider n agents where the first $n/2$ have $v_i = 1$ and $r_i = 1$ with probability 1. The other $n/2$ agents have $v_i = n/2$ and $r_i = 1$ with probability $1/n$ and $r_i = 0$ otherwise. Assume we can probe $k = n/2$ agents and must select $m = 1$ in the end. If we were to select based on $v_i\mathbb{E}_{r_i\sim R_i}[r_i]$, we would select the first $n/2$ agents and would end up with total welfare 1. By selecting the second half, however, we obtain total welfare $\Theta(n)$.

4.1 The Ex-Ante Relaxation

Our first step will be to approximate the objective $\mathrm{obj}(K, v)$ by the ex-ante relaxation $\mathrm{rel}(K, v)$, which we define next. For ease of presentation, it will be convenient to assume that the R_i distributions are continuous (i.e. the CDF is continuous). With this, we can define the *quantile function* $q_i(\theta)$ and the expectation above the quantile $\phi_i(\theta)$. Let $\mathbb{I}[\cdot]$ denote the indicator function and define:

$$q_i(\theta) = \sup\{x \mid \Pr_{r_i\sim R_i}[r_i \ge x] \ge \theta\},$$

$$\phi_i(\theta) = \mathbb{E}_{r_i\sim R_i}[r_i \cdot \mathbb{I}[r_i \ge q_i(\theta)]].$$

Also define $\mathcal{W}$ as the set of feasible allocations and $\mathcal{W}(K)$ as the set of feasible allocations with support in K:

$$\mathcal{W} = \{w \in [0, 1]^n \mid \|w\|_1 \le m\},$$

$$\mathcal{W}(K) = \{w \in \mathcal{W} \mid w_i = 0, \forall i \notin K\}.$$

With that we are ready to define the ex-ante relaxation as:

$$\mathrm{rel}(K, v) = \max_{w\in\mathcal{W}(K)}\left[\sum_{i\in K} v_i \cdot \phi_i(w_i)\right].$$

Next we argue it is a good approximation of the objective:

Lemma 2. *For any $v \in \mathbb{R}_+^n$ and $K \subseteq [n]$,*

$$\mathrm{obj}(K, v) \le \mathrm{rel}(K, v) \le \frac{m!}{m! - e^{-m}m^m}\cdot\mathrm{obj}(K, v) \le \frac{e}{e-1}\cdot\mathrm{obj}(K, v). \tag{2}$$

The proof of Lemma 2 is based on the following result by [7, 19]:

Lemma 3 ([7, 19]). *Let m be an integer and S be a random subset of $[n]$ such that $\mathbb{E}[|S|] \le m$. Then there exists a random subset $\pi(S) \subseteq S$ such that $|\pi(S)| \le m$ almost surely and for each $i \in [n]$ and $\Pr[i \in \pi(S)] \ge \alpha\Pr[i \in S]$ for $\alpha = 1 - e^{-m}m^m/m! \ge 1 - 1/e$.*

Proof of Lemma 2. We first show $\mathrm{rel}(K, v) \ge \mathrm{obj}(K, v)$. Let

$$M^K(v, r) = \operatorname*{argmax}_{M\subseteq K:|M|\le m}\sum_{j\in M} v_j r_j.$$

Throughout the paper we assume argmax gives an arbitrary but consistent single maximizer, instead of the set of all maximizers. Observe that

$$\mathrm{obj}(K, v) = \sum_{i\in[n]}\mathbb{E}\left[v_i r_i \cdot \mathbb{I}\left[i \in M^K(v, r)\right]\right]$$

Define probabilities $p_i^K = \Pr\left[i \in M^K(v, r)\right]$ and observe that $\sum_i p_i^K = \mathbb{E}[|M^K(v, r)|] \le m$. Furthermore, $\mathbb{E}\left[r_i \cdot \mathbb{I}[i \in M^K(v, r)]\right] \le \phi_i(p_i^K)$. This is because $\phi_i(p_i^K) = \sup\{\mathbb{E}[r_i \cdot \mathbb{I}[\mathcal{E}]] \mid \Pr[\mathcal{E}] = p_i^K\} \ge \mathbb{E}[r_i \cdot \mathbb{I}[i \in M^K(v, r)]]$, since $\Pr[i \in M^K(v, r)] = p_i^K$. Therefore:

$$\mathrm{obj}(K, v) \le \sum_{i\in[n]} v_i \cdot \phi_i(p_i^K) \le \mathrm{rel}(K, v).$$

For the second inequality, let $w \in \mathcal{W}(K)$ be a vector such that $\mathrm{rel}(K, v) = \sum_{i\in K} v_i\phi_i(w_i)$ and consider the random set:

$$S_r = \{i \in K \mid r_i \ge q_i(w_i)\}.$$

Observe that $\mathbb{E}[|S_r|] = \sum_{i\in K}\Pr[i \in S_r] = \sum_{i\in K} w_i \le m$. Applying Lemma 3 with $\alpha = 1 - e^{-m}m^m/m!$ we have:

$$\begin{aligned}\mathrm{rel}(K, v) &= \sum_{i\in K} v_i\phi_i(w) = \sum_{i\in K} v_i\mathbb{E}[r_i\mathbb{I}[i \in S_r]] \\ &= \mathbb{E}\left[\sum_{i\in S_r} v_i r_i\right] \le \frac{1}{\alpha}\mathbb{E}_r\left[\sum_{i\in\pi(S_r)} v_i r_i\right].\end{aligned}$$

Since $|\pi(S_r)| \le m$, we clearly have:

$$\mathrm{rel}(K, v) \le \frac{1}{\alpha}\mathbb{E}_r\left[\max_{M\subseteq K;|M|\le m}\sum_{i\in M} v_i r_i\right] = \frac{1}{\alpha}\mathrm{obj}(K, v). \qquad \square$$

Lemma 2 implies that if we obtain a β-approximation with respect to the ex-ante relaxation in the following sense:

$$\mathrm{rel}(f(v), v) \ge \beta\max_{K\in\mathcal{F}}\mathrm{rel}(K, v),$$

then this implies a (β/α)-approximation with respect to the original objective, since:

$$\mathrm{obj}(f(v), v) \ge \frac{1}{\alpha}\mathrm{rel}(f(v), v) \ge \frac{\beta}{\alpha}\max_{K\in\mathcal{F}}\mathrm{rel}(K, v) \ge \frac{\beta}{\alpha}\max_{K\in\mathcal{F}}\mathrm{obj}(K, v).$$

4.2 A Constant-Factor Eligibility Mechanism

Now we are ready to present our constant-factor eligibility mechanism. The approximation factor of this mechanism is $2e/(e-1)$ when $m = 1$, and goes to 2 as m goes to infinity, roughly at a rate of $2 + O(1/\sqrt{m})$. The mechanism f itself is simple: sort all agents by $v_i \cdot \phi_i(m/k)$, and choose the top k agents (ties are broken consistently). Formally,

$$f(v) = \operatorname*{argmax}_{K\subseteq[n]:|K|\le k}\sum_{i\in K} v_i \cdot \phi_i(m/k).$$

The rest of the subsection is devoted to the analysis of the mechanism.

Theorem 3. *For any n, k, m and distributions of $\{r_i\}$, there exists an eligibility mechanism that achieves an approximation factor of $2m!/(m! - e^{-m}) = 2 + O(1/\sqrt{m})$.*

Proof. Let f be the Top-k mechanism stated above. First observe that f is in fact an eligibility mechanism. In particular, it is easy to check f satisfies the condition in Theorem 2. Now consider the approximation factor. Given Lemma 2, we only need to show that for any $v \in \mathbb{R}_+^n$,

$$\mathrm{rel}(f(v), v) \geq \frac{1}{2} \max_{K \subseteq [n]:|K| \leq k} \mathrm{rel}(K, v).$$

Let $K^* = \mathrm{argmax}_{K \subseteq [n]:|K| \leq k} \mathrm{rel}(K, v)$, and $w^* = \mathrm{argmax}_{w \in \mathcal{W}(K^*)} \mathrm{rel}(w, v)$. The plan is to explicitly construct $w \in \mathcal{W}(f(v))$, such that

$$\mathrm{rel}(f(v), v) \geq \sum_i v_i \phi_i(w_i) \geq \frac{1}{2} \sum_i v_i \phi_i(w_i^*) = \frac{1}{2} \mathrm{rel}(K^*, v). \quad (3)$$

Step 1: Construct an intermediary vector $\hat{w} \in \mathcal{W}(K^*)$ such that $\sum_i v_i \phi_i(\hat{w}_i) \geq \frac{1}{2} \sum_i v_i \phi_i(w_i^*)$ and all the non-zero components of $\hat{w}_i$ are at least m/k. To do so, we split K^* in two disjoint sets:

$$A = \{i \in K^* \mid w_i^* \geq m/k\}, \qquad B = \{i \in K^* \mid w_i^* < m/k\}.$$

If $\sum_{i \in A} v_i \phi_i(w_i^*) \geq \frac{1}{2} \sum_{i \in K^*} v_i \phi_i(w_i^*)$, then set $\hat{w}_i = w_i^*$ for $i \in A$ and $\hat{w}_i = 0$ for $i \notin A$. It is straightforward to observe $\hat{w}_i$ satisfies the desired properties.

Otherwise, it must be the case that $\sum_{i \in B} v_i \phi_i(w_i^*) \geq \frac{1}{2} \sum_{i \in K^*} v_i \phi_i(w_i^*)$. In such case, set $\hat{w}_i = m/k$ if $i \in B$ and $\hat{w}_i = 0$ if $i \notin B$. First we observe that $\hat{w} \in \mathcal{W}(K^*)$ since:

$$\sum_i \hat{w}_i \leq |B| \frac{m}{k} \leq k \frac{m}{k} = m.$$

Then observe that that $\hat{w}_i \geq w_i^*$ for $i \in B$, and then:

$$\sum_{i \in B} v_i \phi_i(\hat{w}_i) \geq \sum_{i \in B} v_i \phi_i(w_i^*) \geq \frac{1}{2} \sum_{i \in K^*} v_i \phi_i(w_i^*).$$

Step 2: Construct final vector w satisfying equation (3). Define $\hat{K} = \{i \mid \hat{w}_i > 0\}$ and define a vector w such that $w_i = \hat{w}_i$ for $i \in f(v) \cap \hat{K}$ and $w_i = c$ for $i \in f(v) \setminus \hat{K}$ for some constant c such that $\sum w_i = m$. Since $|f(v)| = m$ and the weights in $f(v) \cap \hat{K}$ are at least m/k, the weights in $f(v) \setminus \hat{K}$ (which are all equal to c) must be at most m/k. Now, define

$$\Phi = \max_{i \in \hat{K} \setminus f(v)} v_i \frac{\phi_i(\hat{w}_i)}{\hat{w}_i}.$$

The core of the proof is encapsulated in the following chain of inequalities where we bound $\phi_i(w_i)$ for each $i \in f(v) \setminus \hat{K}$:

$$\begin{aligned} v_i \frac{\phi_i(w_i)}{w_i} &\geq v_i \frac{\phi_i(m/k)}{m/k} \geq \max_{[n] \setminus f(v)} v_i \frac{\phi_i(m/k)}{m/k} \\ &\geq \max_{\hat{K} \setminus f(v)} v_i \frac{\phi_i(m/k)}{m/k} \geq \max_{\hat{K} \setminus f(v)} v_i \frac{\phi_i(w_i)}{w_i} = \Phi \end{aligned} \quad (4)$$

where the first inequality follows from the fact that $t \mapsto \phi_i(t)/t$ are non-increasing and $w_i \leq m/k$; the second inequality follows from the definition of $f(v)$ and the fourth inequality follows from monotonicity and $w_i \geq m/k$ for $i \in \hat{K}$.

Finally, we put it all together to prove equation (3):

$$\begin{aligned} \sum_{i \in f(v)} v_i \phi_i(w_i) &= \sum_{i \in f(v) \cap \hat{K}} v_i \phi_i(w_i) + \sum_{i \in f(v) \setminus \hat{K}} v_i \phi_i(w_i) \\ &\geq \sum_{i \in f(v) \cap \hat{K}} v_i \phi_i(\hat{w}_i) + \sum_{i \in f(v) \setminus \hat{K}} w_i \Phi \\ &= \sum_{i \in f(v) \cap \hat{K}} v_i \phi_i(\hat{w}_i) + \sum_{i \in \hat{K} \setminus f(v)} \hat{w}_i \Phi \\ &\geq \sum_{i \in f(v) \cap \hat{K}} v_i \phi_i(\hat{w}_i) + \sum_{i \in \hat{K} \setminus f(v)} v_i \phi_i(\hat{w}_i) \\ &\geq \frac{1}{2} \sum_{i \in K^*} v_i \phi_i(w_i^*) \end{aligned}$$

where the first inequality follows from equation (4). The following equality comes the fact that the total weight of w in $f(v) \setminus \hat{K}$ is equal to the total weight of $\hat{w}$ in $\hat{K} \setminus f(v)$ by construction. Finally, the inequalities in the last line follow from the definition of Φ and from the property of $\hat{w}$ established in Step 1. □

4.3 An Alternative Eligibility Mechanism

In this section, we present another quantile-based eligibility mechanism. The approximation factor of this mechanism is not as good as that of the previous mechanism. However, it will useful as a building block for the multi-stage mechanism developed in the next section. The mechanism is again simple. Let

$$\Theta = \{1/2, 1/4, \ldots, 2^{-\lceil \log(k/m) \rceil}\}.$$

The mechanism first chooses some $\theta \in \Theta$ uniformly at random, and then sort all agents by $v_i \cdot \phi_i(\theta)$, and choose the top k agents (ties are broken consistently). Formally, it allocates according to f^θ where

$$f^\theta(v) = \underset{K \subseteq [n]:|K| \leq \min\{k, m/\theta\}}{\mathrm{argmax}} \sum_{i \in K} v_i \cdot \phi_i(\theta),$$

where θ is drawn from Θ uniformly at random. Notably, f is a randomized probing mechanism, and may sometimes choose strictly fewer than k agents.[6]

Theorem 4. *For any n, k, m and distributions of $\{r_i\}$, there exists an eligibility mechanism that achieves an approximation factor of $O(\log(k/m) + 1)$.*

The proof of Theorem 4 actually gives the following properties, which will be useful in the analysis of our multi-stage mechanism.

Corollary 1. *For any n, k, m and distributions of $\{r_i\}$,*

$$\frac{1}{2|\Theta|} \max_{K \subseteq [n]:|K| \leq k} \mathrm{obj}(K, v) \leq \mathbb{E}_\theta \left[\max_{K \subseteq [n]:|K| \leq \min\{k, m/\theta\}} \sum_{i \in K} v_i \cdot \phi_i(\theta) \right],$$

where θ is a uniformly random element from Θ, as defined above. Moreover, for any $\theta \in \Theta$,

$$\max_{K \subseteq [n]:|K| \leq \min\{k, m/\theta\}} \sum_{i \in K} v_i \cdot \phi_i(\theta) \leq O(1) \cdot \max_{K \subseteq [n]:|K| \leq k} \mathrm{obj}(K, v).$$

[6]Choosing k agents is without loss of generality, but we will see that this way of presentation carries more clarity.

Richer feasibility constraints. We remark that this alternative mechanism works for richer feasibility constraints, such as general matroids (while the constant-factor mechanism does not). In particular, consider the setting where the set chosen by the probing mechanism $f(v)$ must be independent in some given matroid, and the objective is to maximize the total value (scaled by r_i) of agents in an independent set (which also must be a subset of $f(v)$) of another given matroid. The single-stage setting discussed above is a special case of this setting where both the outer matroid and the inner matroid are uniform matroids. By adapting the above analysis, one can show that the alternative mechanism guarantees an approximation ratio of $O(\log(\text{rank}_{\text{outer}}/\text{rank}_{\text{inner}}))$ for general matroids. We omit the details here since the proof is essentially the same, only with additional notation.

5 MULTI-STAGE ELIGIBILITY MECHANISMS

So far we studied settings with a single stage of probing followed by an auction. In practice, sometimes there are multiple stages reducing the set of auction participants to fewer and fewer agents (see the quote in footnote 1, for example). Each state corresponds to applying a more accurate, but also more computationally expensive, ML model. In this section we generalize the problem to multi-stage probing and design a truthful approximation (via eligibility mechanisms) for it.

5.1 Multi-Stage Probing Setting

We have again n agents. Each agent $i \in [n]$ is associated with:

- a private value v_i,
- a state space $\mathcal{S}_i$, an initial state $s_{1,i} \in \mathcal{S}_i$ and a Markov chain $s_{t,i}$ on $\mathcal{S}_i$, and
- a function $r_i : \mathcal{S}_i \to \mathbb{R}_+$ denoting the CTR associated with each state.

Only the value v_i is private information of the agent. The algorithm knows from the beginning the complete description of the Markov chain, the initial state for each agent the and r_i functions.

In a multi-stage probing problem we are given a number of stages T, budgets $k_1 \geq k_2 \geq \ldots \geq k_T$ indicating how many agents can be probed in each stage, and a number m of agents that can be selected in the last stage. For notational simplicity let $k_{T+1} = m$. The timing of the process is as follows:

- Algorithm starts with all agents $K_0 = [n]$ knowing only the initial states $s_{1,i}$.
- Algorithm chooses a set $K_1 \subset K_0$ of size at most k_1 to probe. For each $i \in K_1$ we observe $s_{2,i}$.
- Algorithm chooses a set $K_2 \subset K_1$ of size at most k_2 to probe. For each $i \in K_2$ we observe $s_{3,i}$.
- ...
- Algorithm chooses a set $K_T \subset K_{T-1}$ of size at most k_T to probe. For each $i \in K_T$ we observe $s_{T+1,i}$.
- The final CTRs $r_i(s_{T+1,i})$ are observed and the algorithm can choose a set $M \subseteq K_T$ of m agents maximizing $v_i r_i(s_{T+1,i})$.

A multi-stage probing mechanism $f = (f_t)_{t \in [T]}$ consists of a sequence of functions, where each f_t maps the current set of active candidates K_{t-1}, the current state $s_t = (s_{t,i})_{i \in K_{t-1}}$ and the values $v = (v_1, \ldots, v_n)$ to the set $K_t \subseteq K_{t-1}$ with $|K_t| \leq k_t$.

The objective value achieved by f is:

$$\text{obj}(f, v) = \mathbb{E}\left[\max_{M \subseteq K_T : |M| \leq m} \sum_{i \in M} v_i r_i(s_{T+1,i})\right],$$

where the expectation is over the random set K_T which depends on the Markov chains $s_{t,i}$ and the internal randomness of the mechanism f.

As before, we can compose all the probing stages with the winner determination mechanism by recursively applying equation (1). We say that the multi-stage probing mechanism f is an eligibility mechanism if the final composite mechanism is ex-post truthful. The approximation obtained by an eligibility mechanism f is again the ratio between the objective value of the optimal mechanism (not necessarily an eligibility mechanism) and the objective value of f:

$$\alpha = \sup_{\text{general } f'} \sup_{v \in \mathbb{R}_+^n} \frac{\text{obj}(f', v)}{\text{obj}(f, v)}.$$

5.2 Mechanism Construction

We now apply a recursive version of the construction in Theorem 4. The construction will be parametrized by a vector $\theta = (\theta_1, \theta_2, \ldots, \theta_T)$ which will again correspond to quantiles used in each stage.

To define the probing rule f_t^θ in stage t, it will be useful to first define an approximate estimate $\phi_{t,i}^\theta$ for the value we expect to extract from agent i if we select it in stage t. This proxy will again correspond to the solution of the ex-ante relaxation. We define those backwards. For $t = T$, the notion of value will be exactly the same as in Theorem 4 with the distribution of CTRs being the conditioned on the current state $s_{T,i}$. We will define:

$$q_{T,i}^\theta(s_{T,i}) = \sup\{x \in \mathbb{R} \mid \Pr[r_i(s_{T+1,i}) \geq x \mid s_{T,i}] \geq \theta_T\},$$

$$\phi_{T,i}^\theta(s_{T,i}) = \mathbb{E}[r_i(s_{T+1,i}) \cdot \mathbb{I}\{r_i(s_{T+1,i}) \geq q_{T,i}^\theta(s_{T,i})\} \mid s_{T,i}].$$

For all the other stages, we will compute the estimate $\phi_{t,i}^\theta$ as a function of the estimates $\phi_{t+1,i}^\theta$ of the subsequent stage. For $t = T-1, T-2, \ldots, 1$, we define:

$$q_{t,i}^\theta(s_{t,i}) = \sup\{x \in \mathbb{R} \mid \Pr[\phi_{t+1,i}^\theta(s_{t+1,i}) \geq x \mid s_{t,i}] \geq \theta_t\},$$

$$\phi_{t,i}^\theta(s_{t,i}) = \mathbb{E}[\phi_{t+1,i}^\theta(s_{t+1,i}) \cdot \mathbb{I}\{\phi_{t+1,i}^\theta(s_{t+1,i}) \geq q_{t,i}^\theta(s_{T,i})\} \mid s_{t,i}].$$

With this recursive definition in place, we can now define the probing mechanism as follows. First we sample a vector $\theta = (\theta_1, \ldots, \theta_T)$ where each component θ_t is uniformly at random from

$$\Theta_t = \{2^{-i} \mid i = 1, 2, \ldots, \lceil \log(k_t/k_{t+1}) \rceil\}.$$

Then we probe according to:

$$f_t^\theta(K_{t-1}, s_t, v) = \operatorname*{argmax}_{K_t \subseteq K_{t-1}; |K_t| \leq \ell_t^\theta} \sum_{i \in K_t} v_i \phi_{t,i}^\theta(s_{t,i}),$$

where

$$\ell_t^\theta = \min\left\{k_t, \frac{\ell_{t+1}^\theta}{\theta_t}\right\}, \qquad \ell_{T+1}^\theta = m.$$

5.3 Mechanism Analysis

Ex-post truthfulness. First observe that for each fixed parameter θ, the multiple stages of probing composed with the winner determination rule that chooses the m agents to maximize $r_i(s_{T+1,i})v_i$ is ex-post truthful. This is by a recursive application of Theorem 2.

To argue this formally, note that for each $t \in [T]$ the function f_t^θ is an eligibility mechanism (Definition 2). Now we can argue backwards that for every t the composition of $f_t^\theta, f_{t+1}^\theta, \ldots, f_T^\theta$ and the winner determination mechanism is ex-post truthful. Now suppose the above is true for some t. Observe that the composition of f_{t-1}^θ with the composite mechanism starting from t is ex-post truthful, since f_{t-1}^θ is an eligibility mechanism, and the composite mechanism after stage t is ex-post truthful.

Approximation ratio. Finally, we argue about our probing mechanism's approximation ratio. While the approximation ratio is exponential in T, it provides a framework for designing multi-stage mechanisms. In practical applications, it is possible to use heuristics to choose the parameters θ_t instead of choosing them randomly, leading to much better performance.

Theorem 5. *There is an universal constant c such that for any multi-stage setting, the mechanism $\{f_t^\theta\}_{t\in[T]}$ with θ randomly chosen from $\prod_t \Theta_t$ has an approximation factor of*

$$\left(c\left(\frac{\log(k_1/m)}{T}+1\right)\right)^T.$$

A OMITTED PROOFS

Proof of Theorem 2. We first show that the condition is sufficient. Consider any collection of allocation rules $\{x^K\}_K$ that are truthfully implementable. That is, for each K, x^K is non-decreasing in v_i for each $i \in K$. We argue that x^f is non-decreasing in v_i for each $i \in [n]$. Fix some $i \in [n]$, and consider any $v \in \mathbb{R}_+^n$, $r \in [0,1]^n$, and $v_i' \geq v_i$. We only need to show that $x_i^f(v_i', v_{-i}, r) \geq x_i^f(v, r)$ for the composite allocation x^f defined in equation (1). If $i \notin f(v)$, then clearly $x_i^f(v_i', v_{-i}, r) \geq 0 = x_i^f(v, r)$. Suppose $i \in f(v)$. Then $f(v) = f(v_i', v_{-i})$, and

$$\begin{aligned} x_i^f(v, r) &= x_i^{f(v_i', v_{-i})}(v^{f(v_i', v_{-i})}, r^{f(v_i', v_{-i})}) \\ &\leq x_i^{f(v_i', v_{-i})}(v_i', v_{-i}^{f(v_i', v_{-i})}, r^{f(v_i', v_{-i})}) = x_i^f(v_i', v_{-i}, r). \end{aligned}$$

Now we show that the condition is necessary. In particular, whenever the condition does not hold, we construct truthful allocation rules $\{x^K\}_K$ such that x^f is not non-decreasing in some v_i. We consider two cases separately:

- There exists i, v and $v_i' \geq v_i$ where $i \in f(v)$ and $i \notin f(v_i', v_{-i})$.
- There exists i, v and v_i', where $i \in f(v)$, $i \in f(v_i', v_{-i})$, and $f(v) \neq f(v_i', v_{-i})$.

In the first case, consider constant allocation rules $\{x^K\}_K$ (which do not depend on r) where $x_i^K(v') = 1$ whenever $i \in K$ (all other components can be chosen arbitrarily). Then we have $x_i^f(v) = x_i^{f(v)}(v) = 1$ and $x_i^f(v_i', v_{-i}) = x_i^{f(v_i', v_{-i})}(v_i', v_{-i}) = 0$, which means x_i^f is not non-decreasing in i's value. In the second case, without loss of generality suppose $v_i \leq v_i'$, and consider any $j \in f(v_i', v_{-i}) \setminus f(v)$. Consider constant allocation rules $\{x^K\}_K$ (which again do not depend on r) where (1) $x_i^K(v') = 0$ whenever $j \in K$, and (2) whenever $j \notin K$ and $i \in K$, $x_i^K(v') = 1$ (again all other components can be chosen arbitrarily). Then we have $x_i^f(v) = x_i^{f(v)}(v) = 1$ (because $j \notin f(v)$), and $x_i^f(v_i', v_{-i}) = x_i^{f(v_i', v_{-i})}(v_i', v_{-i}) = 0$ (because $j \in f(v_i', v_{-i})$), which means x_i^f is not non-decreasing in i's value. This concludes the proof. □

Proof of Lemma 1. It is known that the following greedy procedure implements the mechanism f (see, e.g., [21]): sort all agents in decreasing order of $s_i(v_i)$. Start with an empty solution, and visit the agents one by one. Upon visiting each agent, add the agent to the solution if the resulting solution is still feasible. Formally, assuming without loss of generality $s_i(v_i) \geq s_{i+1}(v_{i+1})$ for each $i \in [n-1]$, the greedy procedure constructs the following sequence of "prefix" solutions: $S_0 = \emptyset$, and for each $i \in [n]$, $S_i = S_{i-1} \cup \{i\}$ if $S_{i-1} \cup \{i\} \in \mathcal{F}$, and $S_i = S_{i-1}$ otherwise. The output of the greedy procedure is then S_n.

Now consider any $v \in \mathbb{R}_+$, $i \in [n]$, and $v_i' \geq v_i$. We need to show that if i is in the output of the greedy procedure on v, then it is also in the output of the greedy procedure on (v_i', v_{-i}). Let j be the smallest integer such that $s_j(v_j) < s_i(v_i')$[7], and consider the following sequence of sets obtained by running the greedy procedure on (v_i', v_{-i}): for each $i' \in [n]$,

$$T_{i'} = \begin{cases} S_{i'}, & \text{if } i' \leq j \\ T_{i'-1} \cup \{i\}, & \text{if } i' = j \\ T_{i'-1} \cup \{i'-1\}, & \text{if } j < i' \leq i \text{ and } T_{i'-1} \cup \{i'-1\} \in \mathcal{F} \\ T_{i'-1}, & \text{if } j < i' \leq i \text{ and } T_{i'-1} \cup \{i'-1\} \notin \mathcal{F} \\ T_{i'-1} \cup \{i'\}, & \text{if } i' > i \text{ and } T_{i'-1} \cup \{i'\} \in \mathcal{F} \\ T_{i'-1}, & \text{if } i' > i \text{ and } T_{i'-1} \cup \{i'\} \notin \mathcal{F}. \end{cases}$$

We argue that $S_n = T_n$. First observe that $S_{j-1} = T_{j-1}$, since the ways they are obtained are precisely the same. Then observe that for any $i' \in \{j+1, j+2, \ldots, i\}$, $T_{i'} = S_{i'-1} \cup \{i\}$. This is true, because otherwise, we can choose the smallest i' between $j+1$ and i such that $T_{i'} \neq S_{i'-1} \cup \{i\}$. The only way this can happen is $T_{i'} = T_{i'-1}$ and $S_{i'-1} = S_{i'-2} \cup \{i'-1\}$. The first condition (together with the minimality of i') implies $T_{i'-1} \cup \{i'-1\} = S_{i'-2} \cup \{i'-1, i\} \notin \mathcal{F}$, but the second condition implies $S_{i'-2} \cup \{i'-1\} \in \mathcal{F}$, which means $S_{i'-2} \cup \{i'-1, i\} \subseteq S_i \in \mathcal{F}$. This contradicts the downward-closedness of $\mathcal{F}$. Now we have $T_i = S_{i-1} \cup \{i\} = S_i$, which means for any $i' > i$, $T_{i'} = S_{i'}$, and in particular, $T_n = S_n$. This finishes the proof. □

Proof of Theorem 4. The mechanism of probing according to f^θ for a random $\theta \in \Theta$ is an eligibility mechanism, since it is a distribution over mechanisms satisfying the condition in Theorem 2. For the approximation guarantee we will again use Lemma 2. By this lemma, it is enough to show that the following inequality is satisfied by the ex-ante relaxation rel (defined in equation (2)):

$$\max_{K \subseteq [n] : |K| \leq k} \operatorname{rel}(K, v) \leq O(\log(k/m) + 1) \cdot \mathbb{E}_\theta[\operatorname{rel}(f^\theta(v), v)]. \quad (5)$$

Our proof strategy is analogous to the one used in Theorem 3. We initially define $K^* \subseteq [n]$ and $w^* \in \mathcal{W}(K^*)$ such that:

$$K^* = \operatorname*{argmax}_{K \subseteq [n] : |K| \leq k} \operatorname{rel}(K, v), \qquad \operatorname{rel}(K^*, v) = \sum_{i \in K^*} v_i \phi_i(w_i^*).$$

We now decompose w^* into $|\Theta|$ vectors w^θ defined as:

$$w_i^\theta = \begin{cases} w_i^* \cdot \mathbb{I}[\theta < w_i^* \leq 2\theta], & \text{for } \theta \in \Theta \setminus \{2^{-\lceil \log(k/m) \rceil}\} \\ w_i^* \cdot \mathbb{I}[w_i^* \leq 2\theta], & \text{for } \theta = 2^{-\lceil \log(k/m) \rceil} \end{cases}.$$

We will show that:

$$\operatorname{rel}(f^\theta(v), v) \geq \frac{1}{2} \sum_{i \in K^*} v_i \phi_i(w_i^\theta) \quad (6)$$

Taking expectations over equation (6) we obtain equation (5). To show that equation (6) holds, define for each $\theta \in \Theta$ the vector $\tilde{w}^\theta$ as $\tilde{w}_i^\theta = \theta \cdot \mathbb{I}[i \in f^\theta(v)]$. Observe that $\tilde{w}^\theta \in \mathcal{W}(f^\theta(v))$ since $|f^\theta(v)| \leq m/\theta$. Finally, observe that:

$$\begin{aligned} rel(f^\theta(v), v) &= \max_{w \in \mathcal{W}(f^\theta(v))} \sum_i v_i \phi_i(w_i) \\ &\geq \sum_i v_i \phi_i(\tilde{w}_i^\theta) = \sum_{i \in f^\theta(v)} v_i \phi_i(\theta). \end{aligned}$$

Now, let $K^\theta = \{i \in K^* \mid w_i^\theta > 0\}$ be the support of w^θ. Since $w_i^\theta > \theta$ for $i \in K^\theta$ and $\sum_i w_i^\theta \leq \sum_i w_i^* \leq m$ we must have $|K^\theta| \leq |f^\theta(v)|$.

[7] For simplicity we assume all scores are different, but one may check the proof also works with identical scores if ties are broken properly.

Since we are choosing the elements in $f^\theta(v)$ in order of $v_i\phi_i(\theta)$ we must have:

$$\sum_{i\in f^\theta(v)} v_i\phi_i(\theta) \geq \sum_{i\in K^\theta} v_i\phi_i(\theta) \geq \frac{1}{2}\sum_{i\in K^\theta} v_i\phi_i(w_i^\theta),$$

where the last inequality follows from the fact that $w_i^\theta \leq 2\theta$ for $i \in K^\theta$ and the fact that the function $\phi_i(\cdot)$ is concave. The last two display equations together lead to equation (6), concluding the proof. □

Proof of Theorem 5. We will analyze a recursive version of the objective. Let $\mathrm{obj}_t^*(K_{t-1}, s_t, v)$ be the expected objective from executing the optimal mechanism f^* (not necessarily an eligibility mechanism) starting from stage t with an initial set of active candidates K_{t-1} and state s_t. The expectation is taken over the evolution of the Markov chain. It can be defined recursively as:

$$\mathrm{obj}_t^*(K_{t-1}, s_t, v) = \mathbb{E}[\mathrm{obj}_{t+1}^*(f_t^*(K_{t-1}, s_t, v), s_{t+1}, v) \mid s_t],$$

$$\mathrm{obj}_{T+1}^*(K_T, s_{T+1}, v) = \mathbb{E}\left[\max_{M\subseteq K_T; |M|=m} \sum_{i\in M} r_i(s_{T,i})v_i\right].$$

Our first step will be to upper bound obj_t^* using a function of $\phi_{t,i}^\theta$. In particular, we will show that for each t, K_{t-1} and s_t we have:

$$\mathbb{E}_\theta\left[\max_{K_t\subseteq K_{t-1}; |K_t|\leq \ell_t^\theta} \sum_{i\in K_t} v_i\phi_{t,i}^\theta(s_{t,i})\right] \geq \left(\prod_{t\leq t'\leq T}\frac{1}{2|\Theta_{t'}|}\right)\mathrm{obj}_t^*(K_{t-1}, s_t, v). \tag{7}$$

The inequality holds for $t = T$ by Corollary 1. We can show for the remaining values of t by backward induction. First, fix $\theta_{t+1}, \ldots, \theta_T$ and $s_{t,i}$ and apply Corollary 1, which gives:

$$\begin{aligned}
&\mathbb{E}_{\theta_t}\left[\max_{K_t\subseteq K_{t-1}; |K_t|\leq \ell_t^\theta} \sum_{i\in K_t} v_i\phi_{t,i}^\theta(s_{t,i})\right] \\
&\geq \frac{1}{2|\Theta_t|}\max_{K_t\subseteq K_{t-1}; |K_t|\leq k_t} \mathbb{E}\left[\max_{K_{t+1}\subseteq K_t; |K_{t+1}|\leq \ell_{t+1}^\theta} \sum_{i\in K_t} v_i\phi_{t+1,i}^\theta(s_{t+1,i}) \middle| s_t\right].
\end{aligned} \tag{8}$$

Here, n, k, $\min\{k, m/\theta\}$ and m in the single-stage setting correspond to $|K_{t-1}|$, k_t, ℓ_t^θ and ℓ_{t+1}^θ respectively. Moreover, for each i, $\phi_{t+1,i}^\theta(s_{t+1,i})$ corresponds to r_i in the single-stage setting, and $\phi_{t,i}^\theta(s_{t,i})$ correspond to $\phi_i(\theta)$.

Now, observe that $K_t^* := f^*(K_{t-1}, s_{t+1}, v)$ is feasible for the maximization problem in the right hand side of equation (8). This in particular implies that:

$$\begin{aligned}
&\mathbb{E}_{\theta_t}\left[\max_{K_t\subseteq K_{t-1}; |K_t|\leq \ell_t^\theta} \sum_{i\in K_t} v_i\phi_{t,i}^\theta(s_{t,i})\right] \\
&\geq \frac{1}{2|\Theta_t|}\mathbb{E}_\theta\left[\max_{K_{t+1}\subseteq K_t^*; |K_{t+1}|\leq \ell_{t+1}^\theta} \sum_{i\in K_t} v_i\phi_{t+1,i}^\theta(s_{t+1,i}) \middle| s_t\right].
\end{aligned}$$

Now we can take expectations over $\theta_{t+1}, \ldots, \theta_T$ and apply the inductive step (equation (7) for $t+1$), obtaining:

$$\begin{aligned}
&\mathbb{E}_{\theta_t}\left[\max_{K_t\subseteq K_{t-1}; |K_t|\leq \ell_t^\theta} \sum_{i\in K_t} v_i\phi_{t,i}^\theta(s_{t,i})\right] \\
&\geq \frac{1}{2|\Theta_t|}\mathbb{E}\left[\mathrm{obj}_{t+1}^*(f_t(K_{t-1}, s_t, v), s_{t+1}, v) \,\middle|\, s_t\right] \\
&= \frac{1}{2|\Theta_t|}\mathrm{obj}_t^*(K_{t-1}, s_t, v).
\end{aligned}$$

Now that we have established a bound relating the benchmark (i.e., $\phi_{t,i}^\theta$) to the optimal solution, we will establish a bound relating the objective of the f^θ mechanism and the benchmark. For every fixed value of θ we recursively define $\mathrm{obj}_t^\theta(K_t, s_t, v)$ as:

$$\mathrm{obj}_t^\theta(K_{t-1}, s_t, v) = \mathbb{E}[\mathrm{obj}_{t+1}^\theta(f_t^\theta(K_{t-1}, s_t, v), s_{t+1}, v) \mid s_t],$$

$$\mathrm{obj}_{T+1}^\theta(K_T, s_{T+1}, v) = \mathbb{E}\left[\max_{M\subseteq K_T; |M|=m} \sum_{i\in M} r_i(s_{T,i})v_i\right].$$

We will show there is a constant c, such that for any K_{t-1}, s_t and v:

$$\mathrm{obj}_t^\theta(K_{t-1}, s_t, v) \geq c^{T-t+1}\max_{K_t\subseteq K_{t-1}; |K_t|\leq \ell_t^\theta} \sum_{i\in K_t} v_i\phi_{t,i}^\theta(s_{t,i}). \tag{9}$$

For $t = T$, this follows directly from the second part of Corollary 1. We again use backward induction to argue it for the other values of t. First we use the recursive definition of the objective together with the induction hypothesis to see that:

$$\begin{aligned}
&\mathrm{obj}_t^\theta(K_{t-1}, s_t, v) \\
&= \mathbb{E}[\mathrm{obj}_{t+1}^\theta(f_t^\theta(K_{t-1}, s_t, v), s_{t+1}, v) \mid s_{t-1}] \\
&\geq c^{T-t}\cdot\mathbb{E}\left[\max_{K_{t+1}\subseteq f_t^\theta(K_{t-1}, s_t, v); |K_{t+1}|\leq \ell_{t+1}} \sum_{i\in K_{t+1}} v_i\phi_{t+1,i}^\theta(s_{t+1,i}) \middle| s_t\right].
\end{aligned}$$

Then we again apply the second part of Corollary 1 and obtain that:

$$\begin{aligned}
&\mathbb{E}\left[\max_{K_{t+1}\subseteq f_t^\theta(K_{t-1}, s_t, v); |K_{t+1}|\leq \ell_{t+1}} \sum_{i\in K_{t+1}} v_i\phi_{t+1,i}^\theta(s_{t+1,i}) \middle| s_t\right] \\
&\geq c\cdot\sum_{i\in f_t^\theta(K_{t-1}, s_t, v)} v_i\phi_{t,i}^\theta(s_{t,i}) \\
&= c\cdot\max_{K_t\subseteq K_{t-1}; |K_t|\leq \ell_t^\theta} \sum_{i\in K_t} v_i\phi_{t,i}^\theta(s_{t,i}).
\end{aligned}$$

The last two display equations together are a recursive proof of equation (9). To complete the proof, take expectations over θ in equation (9) and combine it with equation (7), both for $t = 1$. Finally, observe that:

$$\begin{aligned}
c^T\prod_t(2|\Theta_t|) &\leq (2c)^T\prod_t\left(1 + \log\frac{k_t}{k_{t+1}}\right) \\
&\leq (2c)^T\left(\frac{\sum_t\left(1 + \log\frac{k_t}{k_{t+1}}\right)}{T}\right)^T = (2c)^T\left(1 + \frac{\log\frac{k_1}{m}}{T}\right)^T
\end{aligned}$$ □

REFERENCES

[1] Marek Adamczyk, Maxim Sviridenko, and Justin Ward. 2016. Submodular stochastic probing on matroids. *Mathematics of Operations Research* 41, 3 (2016), 1022–1038.

[2] Saeed Alaei. 2014. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM J. Comput.* 43, 2 (2014), 930–972.

[3] Curtis Bechtel and Shaddin Dughmi. 2021. Delegated Stochastic Probing. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

[4] Andrei Broder, Peter Ciccolo, Evgeniy Gabrilovich, Vanja Josifovski, Donald Metzler, Lance Riedel, and Jeffrey Yuan. 2009. Online expansion of rare queries for sponsored search. In *Proceedings of the 18th international conference on World wide web*. 511–520.

[5] Andrei Z Broder, Peter Ciccolo, Marcus Fontoura, Evgeniy Gabrilovich, Vanja Josifovski, and Lance Riedel. 2008. Search advertising using web relevance feedback. In *Proceedings of the 17th ACM conference on information and knowledge management*. 1013–1022.

[6] Shuchi Chawla, Jason D Hartline, David L Malec, and Balasubramanian Sivan. 2010. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*. 311–320.

[7] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. 2014. Submodular function maximization via the multilinear relaxation and contention resolution schemes. *SIAM J. Comput.* 43, 6 (2014), 1831–1879.

[8] Wei Chen, Wei Hu, Fu Li, Jian Li, Yu Liu, and Pinyan Lu. 2016. Combinatorial multi-armed bandit with general reward functions. In *Advances in Neural Information Processing Systems*. 1651–1659.

[9] Yejin Choi, Marcus Fontoura, Evgeniy Gabrilovich, Vanja Josifovski, Mauricio Mediano, and Bo Pang. 2010. Using landing pages for sponsored search ad selection. In *Proceedings of the 19th international conference on World wide web*. 251–260.

[10] B. Edelman, M. Ostrovsky, and M. Schwartz. 2007. Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords. *American Economic Review* 97 (2007), 242–259.

[11] Hao Fu, Jian Li, and Pan Xu. 2018. A PTAS for a Class of Stochastic Dynamic Programs. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

[12] Ashish Goel, Sudipto Guha, and Kamesh Munagala. 2006. Asking the right questions: Model-driven optimization using probes. In *Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. 203–212.

[13] Sudipto Guha and Kamesh Munagala. 2012. Adaptive uncertainty resolution in bayesian combinatorial optimization problems. *ACM Transactions on Algorithms (TALG)* 8, 1 (2012), 1–23.

[14] Anupam Gupta and Viswanath Nagarajan. 2013. A stochastic probing problem with applications. In *International Conference on Integer Programming and Combinatorial Optimization*. Springer, 205–216.

[15] Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. 2016. Algorithms and adaptivity gaps for stochastic probing. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 1731–1747.

[16] Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. 2017. Adaptivity gaps for stochastic probing: Submodular and XOS functions. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 1688–1702.

[17] Mohammad Taghi Hajiaghayi, Robert Kleinberg, and Tuomas Sandholm. 2007. Automated online mechanism design and prophet inequalities. In *AAAI*, Vol. 7. 58–65.

[18] Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, et al. 2014. Practical lessons from predicting clicks on ads at facebook. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*. 1–9.

[19] Danish Kashaev and Richard Santiago. 2021. A Simple Optimal Contention Resolution Scheme for Uniform Matroids. *arXiv preprint arXiv:2105.11992* (2021).

[20] Jon Kleinberg and Maithra Raghu. 2018. Team performance with test scores. *ACM Transactions on Economics and Computation (TEAC)* 6, 3-4 (2018), 1–26.

[21] Bernhard Korte and Dirk Hausmann. 1978. An analysis of the greedy heuristic for independence systems. In *Annals of Discrete Mathematics*. Vol. 2. Elsevier, 65–74.

[22] Aranyak Mehta, Uri Nadav, Alexandros Psomas, and Aviad Rubinstein. 2020. Hitting the High Notes: Subset Selection for Maximizing Expected Order Statistics. In *Advances in Neural Information Processing Systems*. 15800–15810.

[23] Roger B Myerson. 1981. Optimal auction design. *Mathematics of operations research* 6, 1 (1981), 58–73.

[24] Ester Samuel-Cahn. 1984. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability* (1984), 1213–1216.

[25] Danny Segev and Sahil Singla. 2021. Efficient approximation schemes for stochastic probing and prophet problems. In *Proceedings of the 22nd ACM Conference on Economics and Computation*. 793–794.

[26] Sahil Singla. 2018. Combinatorial optimization under uncertainty: probing and stopping-time algorithms. *Unpublished doctoral dissertation, Carnegie Mellon University* (2018).

[27] Sahil Singla. 2018. The price of information in combinatorial optimization. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2523–2532.

[28] H.R. Varian. 2007. Position Auctions. *International Journal of Industrial Organization* 25 (2007), 1163–1178.

[29] Yiqing Wang, Xiangyu Liu, Zhenzhe Zheng, Zhilin Zhang, Miao Xu, Chuan Yu, and Fan Wu. 2021. On Designing a Two-stage Auction for Online Advertising. *arXiv preprint arXiv:2111.05555* (2021).