Oblivious Online Contention Resolution Schemes

Hu Fu [*] Pinyan Lu [*] Zhihao Gavin Tang[*] Abner Turkieltaub[†] Hongxun Wu[‡]
Jinzhao Wu[§] Qianfan Zhang[‡]

August 16, 2021

Abstract

Contention resolution schemes (CRSs) are powerful tools for obtaining "ex post feasible" solutions from candidates that are drawn from "ex ante feasible" distributions. Online contention resolution schemes (OCRSs), the online version, have found myriad applications in Bayesian and stochastic problems, such as prophet inequalities and stochastic probing.

When the ex ante distribution is unknown, it was unknown whether good CRSs/OCRSs exist with no sample (in which case the scheme is *oblivious*) or few samples from the distribution. In this work, we give a simple $\frac{1}{e}$-selectable oblivious single item OCRS by mixing two simple schemes evenly, and show, via a Ramsey theory argument, that it is optimal. On the negative side, we show that no CRS or OCRS with $O(1)$ samples can be $\Omega(1)$-balanced/selectable (i.e., preserve every active candidate with a constant probability) for graphic or transversal matroids.

1 Introduction

Contention Resolution Schemes (CRSs) were introduced by Chekuri et al. [8] as a tool for rounding fractional solutions in submodular function maximization. These schemes allow one to first optimize, under feasibility constraints, a continuous extension of a discrete function and then round the fractional solution to an integral feasible solution. Feldman et al. [16] extended this framework to online settings, and the resulting Online Contention Resolution Schemes (OCRSs) turn out to be powerful tools for a wide range of applications in Bayesian and stochastic online optimization, such as prophet inequalities [23, 15], stochastic probing [20, 21, 17], and posted pricing mechanisms [7].

More concretely, given a feasibility system, an ex ante feasible solution is a distribution $\boldsymbol{x}$ over feasible sets. According to such a distribution, the elements are included in the solution in a correlated manner. In many problems, however, the elements are selectable, or *active*, only *independently* according to the marginal distributions given by $\boldsymbol{x}$. A contention resolution scheme is a procedure indexed by the distribution $\boldsymbol{x}$ that, given the set of active elements, must select an (ex post) feasible subset of active elements, and aims to guarantee that each element, when active, is selected with a constant probability c. For many applications, this guarantees to retain at least a c fraction of the objective, compared with that of the (unattainable) ex ante solution. An online contention resolution scheme sees each element's status (of being active or not) in an online fashion, and must decide whether to select an element upon its arrival.

[*]ITCS, Shanghai University of Finance and Economics.
{fuhu, lu.pinyan, tang.zhihao}@mail.shufe.edu.cn
[†]University of British Columbia. abner7@cs.ubc.ca
[‡]IIIS, Tsinghua University. {wuhx18, zqf18}@mails.tsinghua.edu.cn
[§]CFCS, Computer Science Dept., Peking University. jinzhao.wu@pku.edu.cn

In many Bayesian or stochastic problems, it is interesting to study whether good algorithms exist in the absence of the distributional knowledge, and, when sample access is allowed to make up for this lack of knowledge, how the performance of best algorithms scales with the number of samples. For example, in revenue optimal mechanism design, mechanisms with no knowledge of or only sample access to the type distributions are known as prior-free or prior-independent mechanisms, and has seen a large literature devoted to them [e.g. 18, 22, 11, 9, 19]. As another example, in the single-item prophet inequality problem, a *single* sample suffices for an algorithm to match the performance of an optimal one equipped with full knowledge of the distribution [26]. We investigate in this work similar problems for CRS and OCRS when the feasibility system is given by a matroid. In particular, we give a simple, provably optimal, algorithm for single-item OCRS with no distributional knowledge ($\boldsymbol{x}$), and we show that no good CRS or OCRS with $O(1)$ samples exist for graphic or transversal matroids.

We now introduce more formalism in order to discuss our discoveries.

1.1 Oblivious CRS and OCRS

A CRS/OCRS is defined with respect to of a universe U and a downward-closed feasibility system $\mathcal{F} \subseteq 2^U$. An ex ante feasible solution is a distribution over sets in $\mathcal{F}$, represented by a vector $\boldsymbol{x}$ in the polytope associated with $\mathcal{F}$: $\boldsymbol{x} \in \mathcal{P}_{\mathcal{F}} := \mathrm{Conv}(\{1_S\}_{S\in\mathcal{F}})$, where $\mathrm{Conv}(\cdot)$ is the operation of taking the convex hull, and 1_S is the indicator vector for a set S. A convex decomposition of $\boldsymbol{x}$ using the vertices of $\mathcal{P}_{\mathcal{F}}$ gives rise to such a distribution, and x_i is the marginal probability with which element i is included when one draws from this distribution. In many applications, either one is prevented from sampling from this distribution, or such a sample does not serve as a good solution; rather, one observes that each element i is selectable, or *active*, *independently*, with probability x_i. A good CRS maps the set of active elements to a feasible subset, guaranteeing that each element, when active, is kept with a constant probability.

Definition 1.1 ([8]). *A* Contention Resolution Scheme *(CRS) π for $\mathcal{F}$ is a procedure indexed by ex ante feasible solutions $\boldsymbol{x} \in \mathcal{P}_{\mathcal{F}}$. For every $\boldsymbol{x} \in \mathcal{P}_{\mathcal{F}}$, for each $S \subseteq U$, $\pi_{\boldsymbol{x}}(\cdot)$ returns a random set $\pi_{\boldsymbol{x}}(S)$ such that, with probability 1, $\pi_{\boldsymbol{x}}(S) \in \mathcal{F}$ and $\pi_{\boldsymbol{x}}(S) \subseteq S$. A CRS π is* c-balanced, *for $c \in [0,1]$, if for each $\boldsymbol{x} \in \mathcal{P}_{\mathcal{F}}$, for each $i \in U$, $i \in \pi_{\boldsymbol{x}}(R(\boldsymbol{x}))$ with probability at least cx_i, where $R(\boldsymbol{x})$ is a random subset of U such that each element j is in $R(\boldsymbol{x})$ independently with probability x_j.*

A CRS π is said to be deterministic *if $\pi_{\boldsymbol{x}}(\cdot)$ is deterministic for each $\boldsymbol{x}$. It is* oblivious *if, for every $S \subseteq U$, the distribution of $\pi_{\boldsymbol{x}}(S)$ and the distribution of $\pi_{\mathbf{y}}(S)$ are identical for any two $\boldsymbol{x}, \mathbf{y} \in \mathcal{P}_{\mathcal{F}}$.* [1]

The elements in the random set $R(\boldsymbol{x})$ are said to be *active*.

Definition 1.2 ([16]). *An* Online Contention Resolution Scheme *(OCRS) with respect to $\mathcal{F}$ is a procedure indexed by ex ante feasible solutions $\boldsymbol{x} \in \mathcal{P}_{\mathcal{F}}$ and an arrival order of elements in U. The elements arrive according to the order, and is revealed whether it is* active *and, if so, the OCRS must decide irrevocably whether to accept it in the output T. An OCRS is* c-selectable*, for $c \in [0,1]$, if for any $\boldsymbol{x} \in \mathcal{P}_{\mathcal{M}}$ and any arrival order, if each element i is active independently with probability x_i, it is kept by the algorithm in the output with probability at least cx_i.*

An OCRS is oblivious *if, for any arrival order, at any point in the procedure, given the set of elements that have arrived and those that have been accepted, and a newly arrived active element i, the probability with which the algorithm accepts i is the same regardless of $\boldsymbol{x}$.*

In Definition 1.2, the arrival order may depend on $\boldsymbol{x}$ but is fixed before the procedure without knowledge of the set of active elements. It is said to be from an *offline adversary*. Both more powerful adversaries, such

[1] In [8], Chekuri et al. additionally require an oblivious CRS to be deterministic. We remove this requirement. This distinction is important for all the results in this work.

as online and almighty ones, and weaker adversaries, such as random order ones, have been studied in the literature; we focus on the offline adversary model throughout this work. Without loss of generality, we assume the elements arrive in the order $1, 2, \ldots, n$.

In the literature, especially that on OCRSs, the case where $\mathcal{F}$ is a matroid on U has been studied the most. We focus on this case in this paper, and denote the matroid by $\mathcal{M}$.

As an illustration, consider the simplest OCRS for the rank-1 uniform matroid. Here $\boldsymbol{x}$ satisfies $\sum_i x_i \leq 1$, and an OCRS may select at most one element. It is easy to see that no deterministic oblivious CRS can be $\Omega(1)$-balanced even in this simple setting, as pointed out by Chekuri et al. [8]. Consider the randomized OCRS that accepts an active element with probability $\frac{1}{2}$ whenever nothing has been selected. By the end, the algorithm selects nothing with probability at least $\frac{1}{2}$, and hence at the arrival of each element, with probability at least $\frac{1}{2}$ nothing has been selected; the element, if active, therefore has a chance of at least $\frac{1}{4}$ for being selected. This OCRS is hence $\frac{1}{4}$-selectable. Note that it is oblivious, since nowhere does it make use of the distribution $\boldsymbol{x}$. An OCRS with knowledge of $\boldsymbol{x}$ can fine-tune its probability of accepting an active element i: if nothing has been accepted when i arrives, it may accept i with probability $\frac{1}{2}/(1 - \frac{1}{2}\sum_{j<i} x_j)$. By a simple induction, it can been that each element i is accepted with probability precisely $\frac{1}{2}$ conditioning on its being active, and with probability precisely $1 - \frac{1}{2}\sum_{j<i} x_j$, the algorithm has accepted nothing when i arrives. This OCRS that first appeared in [3] is $\frac{1}{2}$-selectable and is optimal for this setting.

1.2 Main Results

Our first main result is a simple $\frac{1}{e}$-selectable oblivious OCRS for the uniform rank 1 matroid, beating the $\frac{1}{4}$-selectable oblivious OCRS best known so far. We also show that our scheme is the optimal oblivious OCRS in this setting.

To motivate the design of our OCRS, consider the following simple, almost wild thought experiment: when processing element i, the non-oblivious, $\frac{1}{2}$-selectable OCRS above uses information of $\boldsymbol{x}$ only for the elements that come before i; when we do not know $\boldsymbol{x}$, what if we see the arrived elements as a (partial) sample from the underlying distribution, and emulate the optimal non-oblivious OCRS, as if the distribution is just the empirical distribution given by that sample? To substantiate this thought, when the first active element i arrives, since no preceding element is active, we are led to crudely estimate $\sum_{j<i} x_j$ to be 0, and we accept i with probability $\frac{1}{2}$; when the second active element i' arrives, if we did not accept the first active element, then we would estimate $\sum_{j<i'} x_j$ to be 1, and we should accept i' with probability $\frac{1}{2}/(1 - \frac{1}{2}) = 1$. To summarize, the resulting algorithm accepts the first active element with probability $\frac{1}{2}$; otherwise, it accepts the second active element. The scheme is obviously oblivious.

In Section 2, we show that precisely this OCRS is $\frac{1}{e}$-selectable. It may be a surprise that this simple algorithm is in fact optimal. Our proof of optimality uses Ramsey theory, following a strategy introduced by Correa et al. [10].

Going beyond the uniform rank 1 matroids, Chekuri et al. [8], among other constructions, gave an oblivious $\Omega(1)$-balanced CRS for the unsplittable flow problem in trees, which implies an oblivious $\Omega(1)$-balanced CRS for laminar matroids. For the general matroid setting, it was not known prior to this work whether $\Omega(1)$-balanced/selectable oblivious CRS/OCRS exists. We show in Section 3 that no oblivious CRS can be $\Omega(1)$-balanced even in graphic or transversal matroids. This immediately also implies that no oblivious $\Omega(1)$-selectable OCRS exists in these settings. Our proof shows that a good oblivious CRS must be able to distinguish, given the set of active elements, between a uniform distribution and a distribution with a hidden structure buried, which is statistically impossible. In fact, our construction shows that, even if a CRS has access to a constant number of samples of $R(\boldsymbol{x})$, it still cannot be $\Omega(1)$-selectable. We note that these impossibility results are not computational, and are therefore not conditional on complexity assumptions.

1.3 Further Related Works

CRSs were first developed in the submodular optimization literature [8, 6], and are closely related to *correlation gaps* [2]. The application in submodular optimization requires an additional *monotonicity* property on the CRS; both monotone and non-monotone CRSs are interesting objects of study in various settings (e.g. [5]). As noted above, Chekuri et al. [8] defined oblivious CRSs with the additional requirement that they be deterministic, a constraint that we relax in our definition. It is not difficult to see that no oblivious, deterministic $\Omega(1)$-balanced CRS exists even for choosing one element from a set of two. Showing no good randomized oblivious CRS exists for matroids is more challenging.

OCRSs were formally defined by Feldman et al. [16], although a problem that is equivalent (under disguise) had been studied by Alaei [3] for the uniform matroids. Feldman et al. gave a $\frac{1}{4}$-selectable OCRS for matroids. Lee and Singla [24] obtained a $\frac{1}{2}$-selectable matroid OCRS by a reverse reduction to the matroid prophet inequalities, albeit requiring the latter to be competitive against the stronger ex ante optimal. Ezra et al. [15] gave a 0.337-selectable OCRS for bipartite matchings. Adamczyk and Wlodarczyk [1] studied OCRSs when the elements arrive in a uniformly random order, and obtained, among other results, $\frac{1}{p+1}$-selectable schemes for p matroid intersections in this setting. Lee and Singla [24]'s reduction also gave a $(1 - \frac{1}{e})$-selectable matroid OCRS with random arrival for the uniform rank 1 matroid.

Dughmi [12, 13] in a recent series of two works showed that the celebrated matroid secretary problem [4] is equivalent to the existence of good *universal* OCRSs for general matroids. Interestingly, before showing this equivalence, in the first paper [12], Dughmi conjectured oblivious OCRSs to be a stepping stone toward the matroid secretary problem. Even though this was bypassed in the final proof of the equivalence [13], it is suggestive that oblivious OCRSs may find other applications.

2 An Optimal Oblivious Single Item OCRS

In this section we give an optimal oblivious OCRS when at most one element can be accepted, i.e., for the uniform rank 1 matroid. This is known as the single item setting. For a uniform rank 1 matroid $\mathcal{M}$ on the universe $U = [n]$, $\boldsymbol{x} \in \mathcal{P}_{\mathcal{M}}$ means $x_i \geq 0$ for each i, and $\sum_{i=1}^{n} x_i \leq 1$. Feldman et al. [16] defined a general, simple family of OCRSs that they call *greedy*. In the single-item case, only one such algorithm makes sense, and we term it *the* greedy algorithm in this setting:

Definition 2.1 ([16]). *The* Greedy *algorithm accepts with probability* 1 *the first active element.*

Our next algorithm, to be mixed with the Greedy algorithm, is equally simple:

Definition 2.2. *The* Accept Second *algorithm passes the first active element and accepts the second active one.*

Sections 2.1 and 2.2 prove the following two theorems, respectively.

Theorem 2.3. *Running the Greedy algorithm and the Accept Second algorithm each with probability $\frac{1}{2}$ is a $\frac{1}{e}$-selectable oblivious single element OCRS.*

Theorem 2.4. *For any $\epsilon > 0$, no oblivious single-item OCRS is $(\frac{1}{e} + \epsilon)$-selectable.*

2.1 Analysis of Selectability

To prove Theorem 2.3, we need to show that for every $i \in [n]$, $\mathbf{Pr}[i \text{ is accepted} \mid i \text{ is active}] \geq \frac{1}{e}$. Since our algorithm is an even mixture of Greedy and Accept Second,

$$\mathbf{Pr}\left[i \text{ is accepted} \mid i \text{ is active}\right] = \frac{1}{2} \cdot \mathbf{Pr}\left[i \text{ is the first or second active element} \mid i \text{ is active}\right].$$

Conditioning on i being active, the probability that i is the first active element equals $\prod_{j<i}(1-x_j)$, and the probability that i is the second active element equals $\sum_{j<i} x_j \prod_{k<i,k\neq j}(1-x_k)$. Therefore,

$$\mathbf{Pr}\left[i \text{ is accepted} \mid i \text{ is active}\right] = \frac{1}{2}\left[\prod_{j<i}(1-x_j) + \sum_{j<i} x_j \prod_{k<i,k\neq j}(1-x_k)\right].$$

For each $i \geq 0$, define $f_i(\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{j\leq i}(1-x_j) + \sum_{j\leq i} x_j \prod_{k\leq i,k\neq j}(1-x_k)$. We now lower bound the value of $f_i(\boldsymbol{x})$ for all i and all possible choices of $\boldsymbol{x}$.

Lemma 2.5. *$f_n(\boldsymbol{x}) \geq (1-\frac{1}{n})^n + (1-\frac{1}{n})^{n-1}$ for all $\boldsymbol{x} \in \{\boldsymbol{x} \mid \forall i, x_i \geq 0; \sum_{i=1}^n x_i \leq 1\}$, where the equality is achieved when $\boldsymbol{x} = (\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n})$.*

Proof. We prove the lemma by induction. The base case when $n=1$ is trivial. For $n \geq 2$, since the domain space $\{\boldsymbol{x} \mid \forall i, x_i \geq 0; \sum_{i=1}^n x_i \leq 1\}$ is compact and the function f_n is continuous, we have that the minimum of f_n is attained by some $\boldsymbol{x}^* \in \{\boldsymbol{x} \mid \forall i, x_i \geq 0; \sum_{i=1}^n x_i \leq 1\}$.

Let $\boldsymbol{x}^{\text{uniform}} = (\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n})$. We have $f_n(\boldsymbol{x}^{\text{uniform}}) = (1-\frac{1}{n})^n + (1-\frac{1}{n})^{n-1}$.

Next, we prove that $f_n(\boldsymbol{x}^*) \geq f_n(\boldsymbol{x}^{\text{uniform}})$ by contradiction. Otherwise $\boldsymbol{x}^* \neq \boldsymbol{x}^{\text{uniform}}$ and there exists two indices i, j such that $x_i^* \neq x_j^*$. Fix $\boldsymbol{x}^*_{\text{-}\{i,j\}}$ and consider the following two quantities.

$$p_0 = \prod_{k\in[n]\setminus\{i,j\}}(1-x_k); \qquad p_1 = \sum_{k\in[n]\setminus\{i,j\}} x_k \prod_{t\in[n]\setminus\{i,j,k\}}(1-x_t).$$

We rearrange the formula of $f_n(\boldsymbol{x}^*)$ as the following.

$$\begin{aligned} f_n(\boldsymbol{x}^*) &= (1-x_i^*)(1-x_j^*)\cdot p_0 + x_i^*(1-x_j^*)\cdot p_0 + x_j^*(1-x_i^*)\cdot p_0 + (1-x_i^*)(1-x_j^*)\cdot p_1 \\ &= (1-x_i^*x_j^*)\cdot p_0 + (1-x_i^*)(1-x_j^*)\cdot p_1 \\ &= p_0 + p_1 - (x_i^*+x_j^*)\cdot p_1 + x_i^*x_j^*\cdot(p_1-p_0). \end{aligned} \tag{1}$$

We now examine two cases based on the relation between p_0 and p_1. We construct $\boldsymbol{x}'$ by adjusting the two coordinates x_i^* and x_j^* while keeping all other variables the same as $\boldsymbol{x}^*_{\text{-}\{i,j\}}$.

- If $p_0 > p_1$, let $x_i' = x_j' = \frac{x_i^*+x_j^*}{2}$. Since $x_i^* \neq x_j^*$, we have $x_i'x_j' > x_i^*x_j^*$. By equation (1), we have $f_n(\boldsymbol{x}') < f_n(\boldsymbol{x}^*)$, that contradicts the optimality of $\boldsymbol{x}^*$.

- If $p_0 \leq p_1$, let $x_i' = 0$ and $x_j' = x_i^* + x_j^*$. Since $x_i'x_j' = 0$ and by equation (1), $f_n(\boldsymbol{x}') \leq f_n(\boldsymbol{x}^*)$. Observe that when $x_i' = 0$, $\boldsymbol{x}'_{\text{-}i}$ is an $(n-1)$-dimensional vector and $f_{n-1}(\boldsymbol{x}'_{-i}) = f_n(\boldsymbol{x}')$. Applying the inductive hypothesis, we have

$$\left(1-\frac{1}{n-1}\right)^{n-1} + \left(1-\frac{1}{n-1}\right)^{n-2} \leq f_n(\boldsymbol{x}') \leq f_n(\boldsymbol{x}^*) < f_n(\boldsymbol{x}^{\text{uniform}}) = \left(1-\frac{1}{n}\right)^n + \left(1-\frac{1}{n}\right)^{n-1},$$

  which contradicts the fact that $(1-\frac{1}{n})^n + (1-\frac{1}{n})^{n-1}$ is monotonically decreasing in n. □

Finally, since $(1-\frac{1}{n})^n + (1-\frac{1}{n})^{n-1}$ is monotonically decreasing in n, with the limit equal to $\frac{2}{e}$, we conclude that our algorithm is $\frac{1}{e}$-selectable.

2.2 Optimality

In this section, we prove Theorem 2.4. We first provide a road map of our proof.

- We first define a restricted class of algorithms called *counting-based* strategies and prove that no counting-based strategy is strictly better than $\frac{1}{e}$-selectable.
- Next, we prove that, for any oblivious OCRS, there must be a subset of elements on which it behaves like a counting-based strategy. This is the most technical step and makes use of Ramsey theorem.
- Finally, we embed the hard instance into the subset and conclude that the hard instance for counting-based strategies applies to all oblivious algorithms.

We start with defining counting-based strategies.

Definition 2.6. *An OCRS is a* counting-based strategy *if it is fully characterized by an infinite sequence of probabilities $(p_1, p_2, \dots)$ as follows: when the algorithm sees the k-th active item, the algorithm accepts (and stops) with probability p_k.*

For instance, the algorithm in Theorem 2.3 is a counting-based strategy characterized by the sequence $(\frac{1}{2}, 1, 0, \dots)$. We now show that it is optimal among all counting-based strategies.

Lemma 2.7. *For every $\epsilon > 0$, there exists a sufficiently large n, such that for the uniform instance with n items (i.e. $x_i = \frac{1}{n}$ for all $i \in [n]$), any counting-based strategy cannot achieve $\left(\frac{1}{e} + \epsilon\right)$-selectability.*

Proof. To prove this, it suffices to show that $\mathbf{Pr}[n \text{ is selected} \mid n \text{ is active}] \le \frac{1}{e} + \epsilon$.

Let q_k be the probability of having exactly k elements active before n. On the uniform instance,

$$q_k = \binom{n-1}{k} \frac{1}{n^k} \left(1 - \frac{1}{n}\right)^{n-1-k} \le \frac{n^k}{k!} \cdot \frac{1}{n^k} \cdot \left(1 - \frac{1}{n}\right)^{n-1-k} = \frac{1}{k!} \left(1 - \frac{1}{n}\right)^{n-1-k}.$$

When $k \ge 3$, $q_k \le \frac{1}{k!} < \frac{1}{e}$. When $k < 3$, $q_k \le \left(1 - \frac{1}{n}\right)^{n-1-k} \to \frac{1}{e}$ when $n \to \infty$. Hence, there is a sufficiently large n such that $q_k \le \frac{1}{e} + \epsilon$ for all k. Then

$$\begin{aligned}\mathbf{Pr}\left[n \text{ is selected} \mid n \text{ is active}\right] &= \sum_{k=1}^{n} q_{k-1} \prod_{i=1}^{k-1} (1 - p_i) \cdot p_k \le \left(\frac{1}{e} + \epsilon\right) \sum_{k=1}^{n} \prod_{i=1}^{k-1} (1 - p_i) \cdot p_k \\ &= \left(\frac{1}{e} + \epsilon\right) \cdot \left(1 - \prod_{i=1}^{n} (1 - p_i)\right) \le \frac{1}{e} + \epsilon,\end{aligned}$$

which concludes the proof. □

Next, we prove that no oblivious OCRS can do better than counting-based strategies in the worst case even when the total number of items N is known to our algorithm.

Given N, any oblivious OCRS can be characterized by a function $f : 2^{[N]} \to [0, 1]$ that specifies the behaviour of the OCRS. In particular, for each set $T \in [N]$, $f(T)$ represents the probability that we select the $(i = \max T)$-th item at step i given that T is the set of active items so far and we have not selected any item yet.

Intuitively, an oblivious OCRS has no information about the vector $\boldsymbol{x}$ and hence, its decision rule should not depend on the indices of the active items. Indeed, for a counting-based strategy, the corresponding function f only depends on the size $|T|$ of the active set. That is, $f(T) = p_{|T|}$ for all T, where $\{p_i\}$ is the probability sequence of the counting-based strategy.

To formalize the intuition, we prove that for any oblivious OCRS, there exists a subset of items $S \subseteq [N]$ on which the algorithm performs like a counting-based strategy. We define (ϵ, n)-approximate counting-based strategies on S.

Definition 2.8. *An oblivious OCRS A is (ϵ, n)-approximate to a counting-based strategy O on S if $f(T) \in [p_{|T|}, p_{|T|} + \epsilon]$ for all $T \subseteq S$ and $|T| \leq n$, where $(p_1, p_2, \cdots)$ is the probability sequence of O.*

Lemma 2.9. *For any integer n, m and $\epsilon > 0$, there exists a sufficiently large integer N such that for any oblivious OCRS A, there exists a subset $S \subseteq [N]$ of size m such that A is (ϵ, n)-approximate to a counting-based strategy on S.*

We first use the lemma to give a proof of Theorem 2.4, before proving the lemma itself.

Proof of Theorem 2.4 By Lemma 2.9, for any integer n and $\epsilon > 0$, there exists an integer N such that for any oblivious OCRS A, there exists $S \subseteq [N]$ of size n such that A is (ϵ, n)-approximate to a counting-based strategy O on S.

Let $(p_1, p_2, \cdots)$ be the corresponding probability sequence of O. Consider any instance that is supported on S, i.e. $x_i = 0$ for all $i \notin S$. We have that for each $i \in S$,

$$\begin{aligned}
&\mathbf{Pr}\left[A \text{ selects } i \mid i \text{ is active}\right] \\
=& \sum_{T \ni i; T \subseteq S} \mathbf{Pr}\left[T - i \text{ are the active items before } i \text{ and are not selected by } A\right] \cdot f(T) \\
\leq& \sum_{T \ni i; T \subseteq S} \mathbf{Pr}\left[T - i \text{ are the active items before } i \text{ and are not selected by } O\right] \cdot (p_{|T|} + \epsilon) \\
\leq& \ \mathbf{Pr}\left[O \text{ selects } i \mid i \text{ is active}\right] + \epsilon,
\end{aligned}$$

where the first inequality follows from the fact that whenever we have a chance to select an item, O selects it with smaller probability than A (by Definition 2.8).

Therefore, if A is $(\frac{1}{e} + 2\epsilon)$-selectable, O is $(\frac{1}{e} + \epsilon)$-selectable for all instances defined on S. However, this contradicts Lemma 2.7, that states any counting-based strategy is no better than $(\frac{1}{e} + \epsilon)$-selectable for the uniform instance on S, i.e. $x_i = \frac{1}{n}$ for $i \in S$. This concludes the proof of the theorem.

Proof of Lemma 2.9. We use the Hypergraph Ramsey theorem. For completeness, we include a few basic concepts of hypergraphs. A hypergraph is called a k-uniform hypergraph if each of its edges contains k vertices. A complete k-uniform hypergraph is a hypergraph $G = (V, E)$ where E is the set of all size-k subsets of V. A clique in a k-uniform hypergraph is a subset $S \subset V$ such that all size-k subsets of S are in E.

Lemma 2.10 (Hypergraph Ramsey Theorem [25, 14]). *Given any positive integer n_0, k and c, there is an integer n_1 which has the following property: for any complete k-uniform hypergraph with more than n_1 vertices, no matter how we color its edges with c colors, there is always a monochromatic clique of size n_0.*

We now prove Lemma 2.9 by induction. For the base case when $n = 1$, consider a graph G with $N = \left(\lfloor \frac{1}{\epsilon} \rfloor + 1\right) m$ vertices. Given an oblivious OCRS, we color each vertex i with color $\lfloor \frac{1}{\epsilon} f(\{i\}) \rfloor$. Since there are at most $\lfloor \frac{1}{\epsilon} \rfloor + 1$ different colors, by the pigeonhole principle, there are at least m vertices sharing the same color. Let S be the set of these m vertices and c be their color. We claim that A is $(\epsilon, 1)$-approximate to any counting-based strategy on S with $p_1 = c\epsilon$.

Suppose the lemma holds for $n - 1$, we consider the case of n. Applying Lemma 2.10 to $n_0 = m, k = n$ and $c = \lfloor \frac{1}{\epsilon} \rfloor + 1$, let n_1 be the sufficiently large number such that for any complete k-uniform hypergraph with more than n_1 vertices, no matter how we color its edges with c colors, there exists a monochromatic clique of size n_0.

By induction hypothesis, there exists N such that for any oblivious OCRS π, there exists a subset $S_0 \subseteq [N]$ of size n_1 such that π is $(\epsilon, n - 1)$-approximate to a counting-based strategy on S_0. Now we consider a complete n-uniform hypergraph G whose vertices are S_0. For each hyperedge $T \subset S_0$ of size n, we color the edge with color $\lfloor \frac{1}{\epsilon} f(T) \rfloor$. Observe that there are at most $\lfloor \frac{1}{\epsilon} \rfloor + 1$ different colors. By Lemma 2.10, this hypergraph includes a monochromatic clique of size m. We denote the vertex set of this clique by S and let c be its corresponding color. By definition of the hypergraph, we have that $f(T) \in [c\epsilon, (c+1)\epsilon]$ for $T \subset S$ and $|T| = n$. Combining this with the inductive hypothesis that π is $(\epsilon, n-1)$-approximate to a counting-based strategy on $S \subseteq S_0$, we conclude that π is further (ϵ, n)-approximate to a counting-based strategy on S with $p_{|T|} = c\epsilon$.

3 Non-existence of $\Omega(1)$-selectable Oblivious CRSs for Matroids

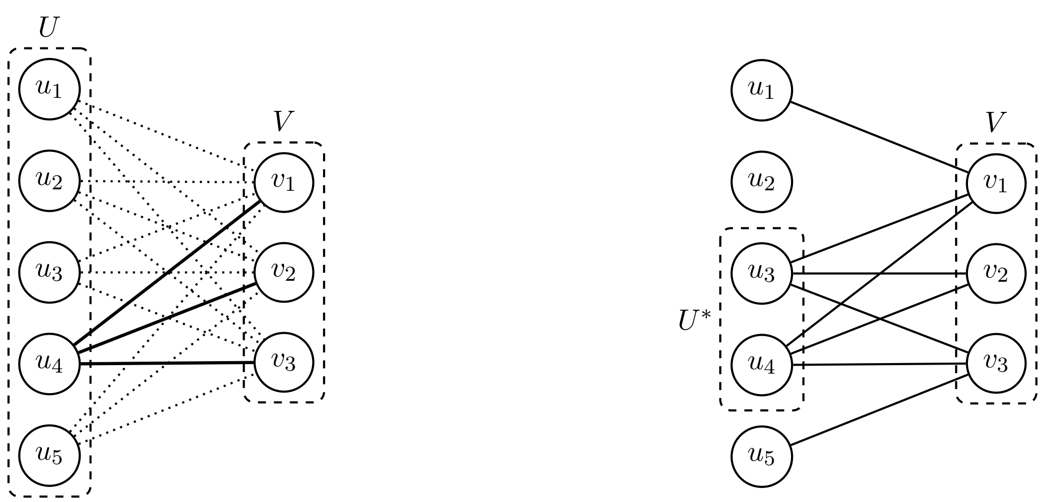
In this section we show that, without knowledge of $\boldsymbol{x}$, $\Omega(1)$-selectable OCRS is impossible for general matroids. In fact, even in the offline setting, $\Omega(1)$-balanced CRS does not exist for graphic or transversal matroids. Furthermore, no such CRS exists even with a constant number of samples. We provide the proofs for graphic matroids and transversal matroids in Section 3.1 and Section 3.2, respectively.

Theorem 3.1. *For any $c \in (0, 1]$, there is no oblivious c-balanced CRS for graphic matroids or transversal matroids. Moreover, the impossibility persists even if the CRS has access to $O(1)$ samples of $R(\boldsymbol{x})$.*

3.1 Graphic Matroids

Recall that a graphic matroid $(E, \mathcal{I})$ is defined on an undirected graph G with edge set E such that $I \subseteq E$ is in $\mathcal{I}$ iff I is a forest in G.

Consider a complete bipartite graph $K_{N,M}$ with bipartition $U = \{u_1, \dots, u_N\}$ and $V = \{v_1, \dots, v_M\}$. Let $\delta(u)$ denote the set of edges incident to a vertex $u \in U \cup V$. Let $\mathcal{P}$ be the polytope of the graphic matroid on it.



(a) The bipartite complete graph $K_{N,M}$. Here $i = 4$, edges adjacent to u_i has probability $x_e^i = 1$ of being active, while other edges each only has probability $1/M$ of being active. Here N should be a large enough number such that $N \gg M^M$.

(b) A realization $R(\boldsymbol{x})$ of this instance. U^* is the set of all vertices on left side of degree M. If N is large enough there will be many vertices happen to be in U^*. These vertices in U^* are indistinguishable to CRS, and u_4 ($i = 4$) is hidden between them.

Figure 1: The hard instance for graphic matroids

We give $N + 1$ points $\boldsymbol{x}^1, \dots, \boldsymbol{x}^N, \boldsymbol{y}$ in the matroid polytope $\mathcal{P}$, and show that, for any $c \in (0, 1]$, there are large enough N and M, such that no oblivious CRS can be c-balanced for the graphic matroid on the complete bipartite graph $K_{N,M}$. The point $\boldsymbol{y}$ is auxiliary in the proof.

For any $i \in [N]$, let $\boldsymbol{x}^i$ be

$$x^i_e := \begin{cases} 1, & \text{if } e \in \delta(u_i); \\ 1/M, & \text{otherwise.} \end{cases}$$

Let $\boldsymbol{y}$ be the vector with weight $\frac{1}{M}$ on all edges, i.e., $y_e = \frac{1}{M}$ for all e. Note that $\boldsymbol{y} = \frac{N}{N+M-1}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}^i$.

We first verify $\boldsymbol{x}^i \in \mathcal{P}$ for each i, which immediately implies $\boldsymbol{y} \in \mathcal{P}$ as well. Let $T_{i,j}$ be the tree whose edge set is $\delta(u_i) \cup \delta(v_j)$. It is easy to verify that $\boldsymbol{x}^i$ is the average of the indicator vectors of $T_{i,j}$ as j ranges from 1 to M; that is, $\boldsymbol{x}^i = \frac{1}{M}\sum_{j=1}^{M}\mathbb{1}_{T_{i,j}}$, where $\mathbb{1}_{T_{i,j}}$ is the indicator vector of $T_{i,j}$. Finally consider the random set U^* defined as follows:

$$U^* = \{u_i \in U \mid \delta(u_i) \subseteq R(\boldsymbol{y})\}.$$

For the sake of a contradiction, suppose there is a c-balanced CRS π for $\mathcal{P}$. We analize the CRS on $R(\boldsymbol{y})$ by considering the expected number of elements accepted from the set of edges incident to U^* (i.e. $\delta(U^*)$). We will first give an upper bound using the feasibility constraints. Then we will provide a lower bound using the assumption that the CRS is c-balanced to get a contradiction. Let $\pi(R(\boldsymbol{y}))$ be the set of accepted elements, then since the output must be an independent set, we have:

$$\mathbf{E}\left[|\pi(R(\boldsymbol{y})) \cap \delta(U^*)|\right] \leq \mathbf{E}\left[\operatorname{rank}(\delta(U^*))\right].$$

And the rank of $\delta(U^*)$ is just $|U^*| + |V| - 1$. Observe that $|U^*|$ follows a binomial distribution with parameters N, M^{-M} (each vertex u_i belongs to U^* with probability M^{-M} independently). So we get the following upper bound:

$$\mathbf{E}\left[|\pi(R(\boldsymbol{y})) \cap \delta(U^*)|\right] \leq \frac{N}{M^M} + M - 1.$$

On the other hand, $\mathbf{E}[|\pi(R(\boldsymbol{y})) \cap \delta(U^*)|]$ can be rewritten as:

$$\sum_{e \in E} \mathbf{Pr}\left[e \in \pi(R(\boldsymbol{y})) \mid e \in \delta(U^*)\right] \mathbf{Pr}\left[e \in \delta(U^*)\right].$$

Take an arbitrary edge $e = (u_i, v_j)$. First note that the probability of $e \in \delta(U^*)$ is just the probability of $u_i \in U^*$, so it is equal to M^{-M}. The crucial observation is that the distribution of $R(\boldsymbol{y})$ conditioned on $(u_i, v_j) \in \delta(U^*)$ is identical to the distribution of $R(x^i)$, so using the assumption that π is c-balanced and that (u_i, v_j) is always active for x^i, we get:

$$\mathbf{Pr}\left[(u_i, v_j) \in \pi(R(\boldsymbol{y})) \mid (u_i, v_j) \in \delta(U^*)\right] = \mathbf{Pr}\left[(u_i, v_j) \in \pi(R(x^i))\right] \geq c.$$

Since there are NM edges in total, we obtain the following lower bound:

$$\mathbf{E}\left[|\pi(R(\boldsymbol{y})) \cap \delta(U^*)|\right] \geq \frac{cN}{M^{M-1}}.$$

By putting together our upper and lower bounds, we get that:

$$c \leq \frac{1}{M} + \frac{M^{M-1}(M-1)}{N},$$

which can be arbitrarily small for $N \gg M^M$ and $M \gg 1$. This finishes the proof in the oblivious case.

For the case in which the CRS has access to a constant number of samples, the idea is almost the same. The only difference is that now we define U^* as the set of vertices u such that $\delta(u)$ is contained in every sample from $R(\boldsymbol{y})$ and in the final realization of $R(\boldsymbol{y})$. If we have s samples, then each vertex has a probability $M^{-(s+1)M}$ of being in U^*. The same arguments hold, we just need $N \gg M^{(s+1)M}$ for the contradiction.

3.2 Transversal Matroids

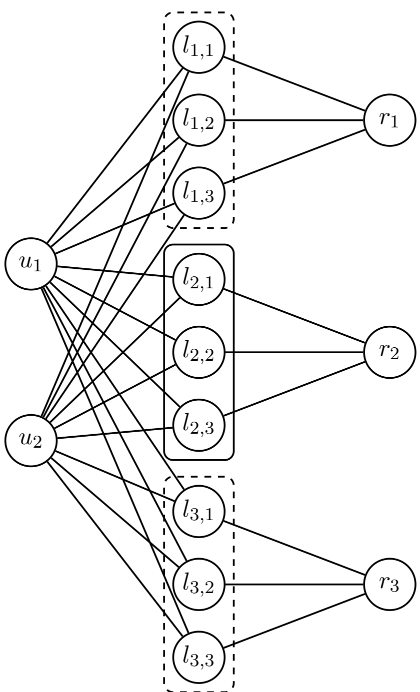
Recall that a transversal matroid $\mathcal{M} = (L, \mathcal{I})$ can be defined by a bipartite graph $G = (L \sqcup R, E)$ such that $I \in \mathcal{I}$ iff there is a matching in G covering I. The proof is very similar to the one provided for graphic matroids.

Consider the transversal matroid defined by the bipartite graph $G = (L \sqcup R, E)$ (see Figure 2a), where

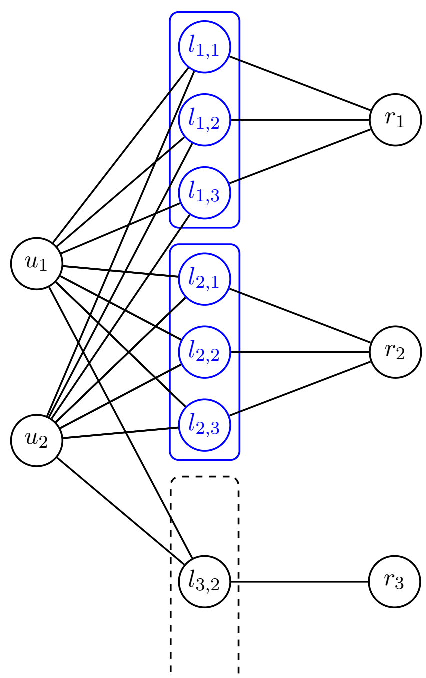
- $L = \{l_{i,j} \mid i \in [N], j \in [M]\}$,
- $R = \{r_i \mid i \in [N]\} \cup \{u_k \mid k \in [M-1]\}$,
- $E = \{(l_{i,j}, r_i) \mid i \in [N], j \in [M]\} \cup \{(l_{i,j}, u_k) \mid i \in [N], j \in [M], k \in [M-1]\}$.

In other words, L contains N sets of M vertices, namely $L = \bigcup_{i \in [N]} L_i$ with $L_i = \{l_{i,j} \mid j \in [M]\}$; for each L_i there is a vertex r_i fully connected to it. Additionally, there is a set of $M-1$ vertices in R fully connected to L. Now, we claim that for any $i' \in [N]$, the following point $\boldsymbol{x}^{i'}$ belongs to the matroid polytope:

$$x^{i'}_{l_{i,j}} \coloneqq \begin{cases} 1, & \text{if } i = i'; \\ 1/M, & \text{otherwise.} \end{cases}$$



(a) The bipartite graph $G = (L \sqcup R, E)$. There are N sets in L with M vertices each. Here $i' = 2$. For each vertex $l_{2,j}$ in the second set, $x_{l_{2,j}} = 1$, while vertices in the other sets each only has $1/M$ probability of being active. The vertices in set $R = \{u_1, \ldots, u_{M-1}\} \cup \{r_1, \ldots, r_N\}$ are drawn on two separated sides.



(b) A realization $R(\boldsymbol{x}^{i'})$ of this instance. A set belongs to U^* if and only if all vertices in it are active. They are marked in blue. As long as N is large enough, there will be many sets in U^* with high probability. These sets are indistinguishable to our CRS. Hence the i'-th set is hidden inside.

Figure 2: The hard instance for transversal matroids

To see that $\boldsymbol{x}^{i'} \in \mathcal{P}_{\mathcal{M}}$, define for each $k \in [M]$ the set $T_{i',k} := \{l_{i',j} \mid j \in [M]\} \cup \{l_{i,k} \mid i \in [N]\}$. It is easy to see that $T_{i',k}$ is independent (we can match each $l_{i,k}$ with r_i and use the vertices u_j to match the $M-1$ remaining vertices). Notice that $\boldsymbol{x}^{i'}$ is the average of the indicator vectors of $T_{i',k}$, that is $\boldsymbol{x}^{i'} = \frac{1}{M}\sum_{k\in[M]} \mathbb{1}_{T_{i',k}}$, so $\boldsymbol{x} \in \mathcal{P}_{\mathcal{M}}$. Again, let y be the vector with weight $\frac{1}{M}$ on all vertices from L. We can see that $y \in \mathcal{P}_{\mathcal{M}}$ since $y \leq \boldsymbol{x}^{i'}$. The rest of proof is almost identical to the previous proof. Instead of U^*, we define now the random set of indices:

$$I^* = \{i \in [N] \mid L_i \subseteq R(y)\},$$

and then we compute bounds for the number of accepted elements from $L^* = \bigcup_{i\in I^*} L_i$. The rank of L^* is $|I^*| + M - 1$: for each $i \in I^*$ we can match one vertex from L_i to r_i, and then we can match $M-1$ more vertices from L^* using the vertices u_t (it is not possible to match more vertices since the rest of the vertices from R are not connected with L^*). Note that the size of I^* also follows a binomial distribution with parameters N and M^{-M}, so we get the exact same upper bound we got before. We also get the same lower bound using equivalent arguments and that concludes the proof.

References

[1] Marek Adamczyk and Michal Wlodarczyk. Random order contention resolution schemes. In Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 790–801. IEEE Computer Society, 2018. 1.3

[2] Shipra Agrawal, Yichuan Ding, Amin Saberi, and Yinyu Ye. Price of correlations in stochastic optimization. *Oper. Res.*, 60(1):150–162, 2012. 1.3

[3] Saeed Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM J. Comput.*, 43(2):930–972, 2014. 1.1, 1.3

[4] Moshe Babaioff, Nicole Immorlica, David Kempe, and Robert Kleinberg. Matroid secretary problems. *J. ACM*, 65(6):35:1–35:26, 2018. URL `https://doi.org/10.1145/3212512`. 1.3

[5] Simon Bruggmann and Rico Zenklusen. An optimal monotone contention resolution scheme for bipartite matchings via a polyhedral viewpoint. *CoRR*, abs/1905.08658, 2019. 1.3

[6] Gruia Călinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a submodular set function subject to a matroid constraint (extended abstract). In *IPCO*, volume 4513 of *Lecture Notes in Computer Science*, pages 182–196. Springer, 2007. 1.3

[7] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In Leonard J. Schulman, editor, *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 311–320. ACM, 2010. 1

[8] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. *SIAM J. Comput.*, 43(6):1831–1879, 2014. 1, 1.1, 1, 1.1, 1.2, 1.3

[9] Richard Cole and Tim Roughgarden. The sample complexity of revenue maximization. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 243–252. ACM, 2014. 1

[10] José R. Correa, Paul Dütting, Felix A. Fischer, and Kevin Schewior. Prophet inequalities for I.I.D. random variables from an unknown distribution. In Anna Karlin, Nicole Immorlica, and Ramesh Johari, editors, *Proceedings of the 2019 ACM Conference on Economics and Computation, EC 2019, Phoenix, AZ, USA, June 24-28, 2019*, pages 3–17. ACM, 2019. 1.2

[11] Peerapong Dhangwatnotai, Tim Roughgarden, and Qiqi Yan. Revenue maximization with a single sample. *Games Econ. Behav.*, 91:318–333, 2015. 1

[12] Shaddin Dughmi. The outer limits of contention resolution on matroids and connections to the secretary problem. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 42:1–42:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 1.3

[13] Shaddin Dughmi. Matroid secretary is equivalent to contention resolution. *CoRR*, abs/2103.04205, 2021. URL `https://arxiv.org/abs/2103.04205`. 1.3

[14] Paul Erdos and Richard Rado. Combinatorial theorems on classifications of subsets of a given set. *Proceedings of the London mathematical Society*, 3(1):417–439, 1952. 2.10

[15] Tomer Ezra, Michal Feldman, Nick Gravin, and Zhihao Gavin Tang. Online stochastic max-weight matching: Prophet inequality for vertex and edge arrival models. In Péter Biró, Jason Hartline, Michael Ostrovsky, and Ariel D. Procaccia, editors, *EC '20: The 21st ACM Conference on Economics and Computation, Virtual Event, Hungary, July 13-17, 2020*, pages 769–787. ACM, 2020. 1, 1.3

[16] Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 1014–1033. SIAM, 2016. 1, 1.2, 1.3, 2, 2.1

[17] Hu Fu, Zhihao Gavin Tang, Hongxun Wu, Jinzhao Wu, and Qianfan Zhang. Random order vertex arrival contention resolution schemes for matching, with applications. In *ICALP*, volume 198 of *LIPIcs*, pages 68:1–68:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. 1

[18] Andrew V. Goldberg, Jason D. Hartline, Anna R. Karlin, Michael E. Saks, and Andrew Wright. Competitive auctions. *Games Econ. Behav.*, 55(2):242–269, 2006. 1

[19] Yannai A. Gonczarowski and S. Matthew Weinberg. The sample complexity of up-to-ϵ multi-dimensional revenue maximization. *J. ACM*, 68(3):15:1–15:28, 2021. 1

[20] Anupam Gupta and Viswanath Nagarajan. A stochastic probing problem with applications. In Michel X. Goemans and José R. Correa, editors, *Integer Programming and Combinatorial Optimization - 16th International Conference, IPCO 2013, Valparaíso, Chile, March 18-20, 2013. Proceedings*, volume 7801 of *Lecture Notes in Computer Science*, pages 205–216. Springer, 2013. 1

[21] Anupam Gupta, Viswanath Nagarajan, and Sahil Singla. Algorithms and adaptivity gaps for stochastic probing. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 1731–1747. SIAM, 2016. 1

[22] Jason D. Hartline and Tim Roughgarden. Optimal mechanism design and money burning. In Cynthia Dwork, editor, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008*, pages 75–84. ACM, 2008. 1

[23] Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities and applications to multi-dimensional mechanism design. *Games Econ. Behav.*, 113:97–115, 2019. 1

[24] Euiwoong Lee and Sahil Singla. Optimal online contention resolution schemes via ex-ante prophet inequalities. In Yossi Azar, Hannah Bast, and Grzegorz Herman, editors, *26th Annual European Symposium on Algorithms, ESA 2018, August 20-22, 2018, Helsinki, Finland*, volume 112 of *LIPIcs*, pages 57:1–57:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. 1.3

[25] FP Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 2(1):264–286, 1930. 2.10

[26] Aviad Rubinstein, Jack Z. Wang, and S. Matthew Weinberg. Optimal single-choice prophet inequalities from samples. In Thomas Vidick, editor, *11th Innovations in Theoretical Computer Science Conference, ITCS 2020, January 12-14, 2020, Seattle, Washington, USA*, volume 151 of *LIPIcs*, pages 60:1–60:10. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 1