Universal Online Contention Resolution with Preselected Order

Junyao Zhao[*]

IRIF, CNRS, Université Paris Cité

junyao-zhao@outlook.com

Abstract

Online contention resolution scheme (OCRS) is a powerful technique for online decision making, which—in the case of matroids—given a matroid and a prior distribution of active elements, selects a subset of active elements that satisfies the matroid constraint in an online fashion. OCRS has been studied mostly for product distributions in the literature. Recently, universal OCRS, that works even for correlated distributions, has gained interest, because it naturally generalizes the classic notion, and its existence in the random-order arrival model turns out to be equivalent to the matroid secretary conjecture. However, currently very little is known about how to design universal OCRSs for any arrival model. In this work, we consider a natural and relatively flexible arrival model, where the OCRS is allowed to preselect (i.e., non-adaptively select) the arrival order of the elements, and within this model, we design simple and optimal universal OCRSs that are computationally efficient. In the course of deriving our OCRSs, we also discover an efficient reduction from universal online contention resolution to the matroid secretary problem for any arrival model, answering a question from Dughmi (2020).

1 Introduction

A contention resolution scheme (CRS) is an algorithm which given a set system $\mathcal{M} \subseteq 2^{[n]}$ and a set of *active* elements $A \subseteq [n]$ sampled from a known prior distribution $\mathcal{D}_A$, selects a subset $X \subseteq A$ that satisfies the feasibility constraint $X \in \mathcal{M}$. The design goal of CRS is to guarantee that every element in $[n]$ is selected with some constant probability α conditioned on it being active. Informally, when such CRS exists for set system $\mathcal{M}$ and prior distribution $\mathcal{D}_A$, we call $\mathcal{D}_A$ an α-*uncontentious* distribution for $\mathcal{M}$. CRS was introduced by Chekuri et al. (2014) and has since been studied for various set systems. In this paper, we focus on a most studied set system in the literature—matroid (see Definition 2.1).

A particular class of CRSs, known as online contention resolution schemes (OCRSs), has found many applications in online decision making, such as prophet inequalities, stochastic probing, and sequential posted-price auctions (e.g., Gupta and Nagarajan (2013); Feldman et al. (2021); Adamczyk and Włodarczyk (2018)). Specifically, an OCRS knows only $\mathcal{M}$ and $\mathcal{D}_A$ but not the set of active elements A at the beginning. Instead, elements in $[n]$ arrive one by one (their order depends on the specific arrival model), and upon the arrival of each element $i \in [n]$, it is revealed whether $i \in A$, and the OCRS must decide immediately and irrevocably whether to include i in its solution set X before the next element arrives.

Many elegant OCRSs with strong guarantees (in various arrival models) have been discovered (e.g., Chekuri et al. (2014); Feldman et al. (2021); Adamczyk and Włodarczyk (2018); Lee and Singla

---

(2018); Fu et al. (2024)), but most of them were established exclusively for *product* distribution $\mathcal{D}_A$ (i.e., each element is active independently with some probability), except for Dughmi et al. (2024) and Gupta et al. (2024), who studied OCRSs for pairwise independent distribution $\mathcal{D}_A$.

Recently, *universal* OCRSs (Dughmi, 2020), that work even for general *correlated* distribution $\mathcal{D}_A$, have started to gain interest. Informally, an OCRS is (α, β)-universal if it guarantees that every element is selected with some constant probability β conditioned on it being active, for any matroid $\mathcal{M}$ and (arbitrarily correlated) α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$. Universal OCRSs naturally generalize classic OCRSs for product distributions. Moreover, Dughmi (2020, 2022) proved that the existence of universal OCRSs in the random-order arrival model is equivalent to the matroid secretary conjecture posed by Babaioff et al. (2007).

Currently, very little is known about how to design universal OCRSs, except that Dughmi (2020) showed that for any arrival model[1], a universal OCRS exists if there is a constant-competitive matroid secretary algorithm for that model. For the free-order arrival model, where the algorithm is allowed to *adaptively* choose the arrival order of the remaining elements after observing any number of elements, Jaillet et al. (2013) designed a constant-competitive matroid secretary algorithm. This algorithm, together with Dughmi's result, implies that there exists a universal OCRS in the free-order model. However, this approach is not known to be computationally efficient, making it difficult to understand the implied universal OCRS for specific problem instances. Indeed, Dughmi's result is information-theoretic, and whether it can be made computationally efficient was left as an open question (Dughmi, 2020, Section 6).

In this work, we strive to design universal OCRSs that are efficiently computable and simple to understand. We focus on a natural and relatively flexible arrival model in which the OCRS is allowed to preselect the arrival order of the elements (i.e., *non-adaptively* choose the arrival order of the elements given $\mathcal{M}$ and $\mathcal{D}_A$). We adopt the term *preselected order* to differentiate from the free-order model. The preselected-order model is a step toward the random-order model, and it has been studied for various online decision problems, including prophet inequalities (Hill, 1983; Agrawal et al., 2020; Liu et al., 2021; Peng and Tang, 2022; Bubna and Chiplunkar, 2023), the multi-choice secretary problem (Hajiaghayi et al., 2022), sequential posted-price auctions (Chawla et al., 2010; Beyhaghi et al., 2018), stochastic probing (Gupta and Nagarajan, 2013) and OCRSs for product distributions (Chekuri et al., 2014). The main contribution of our work is the design and analysis of three different universal OCRSs in the preselected-order model.

1.1 Overview of our universal OCRSs

Now we give an overview of our universal OCRSs. All of our universal OCRSs are generalizations of *ordered* OCRSs for product distributions (Chekuri et al., 2014; Gupta and Nagarajan, 2013) with necessary randomization. Briefly, an ordered OCRS (Gupta and Nagarajan, 2013, Definition 3.2) preselects the arrival order of the elements, and then upon the arrival of each element, it greedily adds the element to the solution set, provided that the element is *selectable* (i.e., if it is active and can be added to the solution set without violating the matroid constraint).

Our first two universal OCRSs (Algorithm 1 and Algorithm 3) generalize ordered OCRSs by *subsampling selectable elements*. Specifically, these two OCRSs first preselect the arrival order (based on their respective criteria) and sample a subset of elements $T \subseteq [n]$. Then, upon the arrival of each element, they add the element to the solution set if it is selectable and belongs to T. Algorithm 1 and Algorithm 3 use different subsampling methods of independent interest— Algorithm 1 includes each element in T independently with a certain probability, while Algorithm 3

---

[1]To be precise, Dughmi's result (Dughmi, 2020, Theorem 4.1) was stated for the random-order arrival model, but the proof of that result applies to any arrival model.

employs a correlated subsampling method and achieves a slightly stronger universality guarantee.

Theorem 1.1 (Restatement of Theorem 3.4 and Theorem 4.4)**.** *For all $\alpha \in [0, 1]$, Algorithm 1 is an $(\alpha, \frac{\alpha^2}{4})$-universal OCRS, and Algorithm 3 is an $(\alpha, \frac{\alpha^2}{2})$-universal OCRS. Both algorithms preselect the arrival order of the elements before any element arrives.*

Instead of subsampling selectable elements, our third universal OCRS (presented in Section 5) *samples the arrival order* of the elements. Specifically, given matroid $\mathcal{M}$ and prior distribution $\mathcal{D}_A$, this OCRS first computes a distribution over permutations by solving a linear program (LP), which is similar to the LP used by Chekuri et al. (2014) to compute offline CRSs for product distributions. Then, it samples the arrival order from this distribution, and upon the arrival of each element, it includes the element in the solution set if the element is selectable. The universality guarantee of this OCRS is nearly optimal[2].

Theorem 1.2 (Informal restatement of Theorem 5.4)**.** *For any $\varepsilon > 0$, there exists a computationally efficient OCRS with preselected order, which is $(\alpha, (1 - \varepsilon)\alpha)$-universal for all $\alpha \in [0, 1]$.*

For comparison, our first two OCRSs have weaker universality guarantees, but they are easier to reason about for specific problem instances. In contrast, our third OCRS provides the strongest universality guarantee, though it is less intuitive because of the use of the LP.

1.2 Universal OCRSs from secretary algorithms

In the course of deriving our third universal OCRS, we discovered an LP-based *efficient reduction* from universal online contention resolution to the matroid secretary problem for any arrival model (see Section 6 for the setup of the matroid secretary problem), thereby answering the aforementioned question from Dughmi (2020).

Theorem 1.3 (Informal restatement of Theorem 6.3)**.** *For any $c, \varepsilon > 0$, for any arrival model, if there is a computationally efficient c-competitive matroid secretary algorithm, then there exists a computationally efficient OCRS that is $(\alpha, (1 - \varepsilon)c \cdot \alpha)$-universal for all $\alpha \in [0, 1]$.*

Briefly, Dughmi's information-theoretic reduction (Dughmi, 2020, Theorem 4.1) was based on the separating hyperplane theorem. Our reduction replaces that with an LP duality argument, and then applies a technique from Lee and Singla (2018) to solve the corresponding LPs efficiently.

2 Preliminaries

2.1 Matroids

We start by introducing essential definitions and properties of *matroids*, and we refer interested readers to Welsh (2010) for a comprehensive treatment of matroid theory. Essentially, a matroid is a set system with some independence structure, which is defined as follows.

Definition 2.1 (matroid)**.** *A set system $\mathcal{M} \subseteq 2^{[n]}$ is a matroid if it satisfies the following properties:*

  i. $\emptyset \in \mathcal{M}$.

  ii. If $X \in \mathcal{M}$, then $Y \in \mathcal{M}$ for all $Y \subseteq X$.

  iii. If $X, Y \in \mathcal{M}$ and $|Y| < |X|$, then there exists $i \in X \setminus Y$ such that $Y \cup \{i\} \in \mathcal{M}$.

---

[2]We note that the ε loss is only due to computation—we will show that (α, α)-universal OCRSs exist.

Given a matroid, we can associate a *(weighted) rank function* with it.

Definition 2.2 (rank)**.** *Given a matroid $\mathcal{M} \subseteq 2^{[n]}$, the rank function $r_{\mathcal{M}} : 2^{[n]} \to \mathbb{Z}_{\geq 0}$ associated with $\mathcal{M}$ is $r_{\mathcal{M}}(X) := \max_{Y \subseteq X} |Y|$ s.t. $Y \in \mathcal{M}$. For convenience, for any sets $X, Y \subseteq [n]$, we denote $r_{\mathcal{M}}(X \mid Y) := r_{\mathcal{M}}(X \cup Y) - r_{\mathcal{M}}(Y)$.*

More generally, for any weight vector $w \in \mathbb{R}^n_{\geq 0}$, the weighted rank function $r_{\mathcal{M}, w} : 2^{[n]} \to \mathbb{R}_{\geq 0}$ is $r_{\mathcal{M}, w}(X) := \max_{Y \subseteq X} \sum_{i \in Y} w_i$ s.t. $Y \in \mathcal{M}$.

Moreover, we can define notions of *span* and *basis* for a matroid.

Definition 2.3 (span)**.** *Given a matroid $\mathcal{M} \subseteq 2^{[n]}$ and its rank function $r_{\mathcal{M}} : 2^{[n]} \to \mathbb{Z}_{\geq 0}$, we say that an element $i \in [n]$ is spanned by a set $X \subseteq [n]$ if $r_{\mathcal{M}}(X \cup \{i\}) = r_{\mathcal{M}}(X)$, and we define the span of X as $\text{span}_{\mathcal{M}}(X) := \{i \in [n] \mid i \text{ is spanned by } X\}$.*

Definition 2.4 (basis)**.** *Given a matroid $\mathcal{M} \subseteq 2^{[n]}$ and a set $X \subseteq [n]$, we say that a subset $Y \subseteq X$ is a basis of X if $Y \in \mathcal{M}$ and for all $i \in X \setminus Y$, $Y \cup \{i\} \notin \mathcal{M}$.*

Furthermore, we define the *restriction* of a matroid to a set of elements.

Definition 2.5 (restriction)**.** *Given a matroid $\mathcal{M} \subseteq 2^{[n]}$ and a set $X \subseteq [n]$, we define the restriction of $\mathcal{M}$ to X as $\mathcal{M}_X := \{Y \subseteq X \mid Y \in \mathcal{M}\}$.*

In Section A, we state several well-known properties of matroids (Lemma A.1).

2.2 Contention resolution schemes

Now we introduce *contention resolution schemes* (CRSs) for matroids. Given a matroid $\mathcal{M} \subseteq 2^{[n]}$ and a random set of *active* elements $A \subseteq [n]$ sampled from a prior distribution[3] $\mathcal{D}_A \in \Delta(2^{[n]})$, a CRS selects a subset of active elements $X \subseteq A$ such that $X \in \mathcal{M}$. Formally, we first let $\Phi_{\mathcal{M}}$ denote the family of maps that take an active set $A \subseteq [n]$ as input and output a subset $X \subseteq A$ such that $X \in \mathcal{M}$, i.e.,

$$\Phi_{\mathcal{M}} := \{\phi : 2^{[n]} \to \mathcal{M} \mid \text{for all } A \subseteq [n], \phi(A) \subseteq A\}.$$

Then, a CRS for $(\mathcal{M}, \mathcal{D}_A)$ is a random map sampled from a distribution $\mathcal{D}_\phi \in \Delta(\Phi_{\mathcal{M}})$ (since the choice of $\mathcal{D}_\phi$ fully determines the CRS, we also use the notation $\mathcal{D}_\phi$ to refer to the CRS itself). Moreover, for $\alpha \in [0, 1]$, we say that a CRS $\mathcal{D}_\phi$ for $(\mathcal{M}, \mathcal{D}_A)$ is α-*balanced* if it satisfies

$$\Pr_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}[i \in \phi(A) \mid i \in A] \geq \alpha \text{ for all } i \in [n] \text{ s.t. } \Pr_{A \sim \mathcal{D}_A}[i \in A] > 0,$$

and we say that $\mathcal{D}_A$ is α-*uncontentious* for $\mathcal{M}$ if there exists an α-balanced CRS for $(\mathcal{M}, \mathcal{D}_A)$.

Furthermore, we use the notation $\mathcal{D}_A^S$ to denote marginal distributions of $\mathcal{D}_A$. That is, for any $S \subseteq [n]$, $\mathcal{D}_A^S$ is defined as follows: $\forall Z \subseteq S$, $\Pr_{A' \sim \mathcal{D}_A^S}[A' = Z] := \Pr_{A \sim \mathcal{D}_A}[A \cap S = Z]$. We note that if $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, then $\mathcal{D}_A^S$ is α-uncontentious for the restriction $\mathcal{M}_S$.

Lemma 2.6. *Given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, for any $S \subseteq [n]$, the marginal distribution $\mathcal{D}_A^S$ is α-uncontentious for the restriction $\mathcal{M}_S$.*

A self-contained proof of Lemma 2.6 is provided in Section B.

---

[3] Throughout the paper, we assume w.l.o.g. that $\mathcal{D}_A$ satisfies that $\forall i \in [n]$, $\Pr_{A \sim \mathcal{D}_A}[i \in A] > 0$. This assumption will simplify the presentation, as it ensures that the probabilities conditioned on the event $i \in A$ are well-defined.

Online contention resolution schemes with preselected order

In this paper, we mostly focus on *online contention resolution schemes* (OCRSs) that first *preselect* the arrival order of the elements and then select a subset of active elements in an *online* fashion. Specifically, an OCRS with preselected order is an algorithm ALG that works in three stages:

(1) Given oracle access[4] to a matroid $\mathcal{M} \subseteq 2^{[n]}$ and a prior distribution $\mathcal{D}_A \in \Delta(2^{[n]})$, ALG first selects a (random) permutation $\pi : [n] \to [n]$ and initializes an empty solution set X_{ALG}.

(2) A random set of active elements $A \subseteq [n]$ is sampled from $\mathcal{D}_A$ and is unknown to ALG.

(3) Then, ALG runs in n steps. At each step $i \in [n]$, it is revealed to ALG whether $\pi(i) \in A$. If $\pi(i)$ is *selectable* (i.e., $\pi(i) \in A$ and $X_{\mathsf{ALG}} \cup \{\pi(i)\} \in \mathcal{M}$), ALG must decide immediately and irrevocably whether to add $\pi(i)$ to X_{ALG}, and it can use randomness to make this decision.

Similar to general CRSs, for $\beta \in [0, 1]$, we say that ALG is *β-balanced* for $(\mathcal{M}, \mathcal{D}_A)$ if it satisfies

$$\Pr\nolimits_{A \sim \mathcal{D}_A, \text{ randomness of } \mathsf{ALG}}[i \in X_{\mathsf{ALG}} \mid i \in A] \geq \beta \text{ for all } i \in [n] \text{ s.t. } \Pr\nolimits_{A \sim \mathcal{D}_A}[i \in A] > 0.$$

Moreover, for $0 \leq \beta \leq \alpha \leq 1$, we say that ALG is *$(\alpha, \beta)$-universal* if for any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any α-uncontentious distribution $\mathcal{D}_A \in \Delta(2^{[n]})$ for $\mathcal{M}$, ALG is β-balanced for $(\mathcal{M}, \mathcal{D}_A)$. Intuitively, universality means that ALG is approximately as balanced as any CRS for any instance $(\mathcal{M}, \mathcal{D}_A)$.

Furthermore, we say ALG is *computationally efficient* if it runs in $O\left(\mathsf{poly}\left(\frac{n}{p_{\min}}\right) \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}})\right)$ time, where $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$, and $t_{\mathcal{D}_A}$, $t_{\mathcal{M}}$ are the time it takes to generate a sample from $\mathcal{D}_A$ and to check whether a set of elements belongs to $\mathcal{M}$ respectively. We note that the runtime must depend on $\frac{1}{p_{\min}}$, if ALG is (α, β)-universal for arbitrary constants $\alpha, \beta \in (0, 1)$.

Example 2.7. *We consider the 1-uniform matroid $\mathcal{M} = \{\emptyset, \{1\}, \{2\}, \ldots, \{n\}\}$. For any constant $\alpha \in (0, 1)$ and any $\delta \in \left(0, \frac{1}{n+1/\alpha-2}\right]$, we consider a family of prior distributions $\{\mathcal{D}_A^{(j)} \mid j \in [n]\}$, where each distribution $\mathcal{D}_A^{(j)}$ is defined as follows:*

$$\Pr[A = \emptyset] = 1 - \delta \cdot \left(n + \frac{1}{\alpha} - 2\right), \quad \Pr[A = [n]] = \delta \cdot \left(\frac{1}{\alpha} - 1\right), \quad \Pr[A = \{i\}] = \delta \text{ for all } i \neq j.$$

Note that $p_{\min} = \Pr[j \in A] = \delta \cdot \left(\frac{1}{\alpha} - 1\right)$ for each prior distribution $\mathcal{D}_A^{(j)}$. Moreover, every prior distribution $\mathcal{D}_A^{(j)}$ is α-uncontentious for matroid $\mathcal{M}$, because the CRS, that selects element j if $A = [n]$ and selects element i if $A = \{i\}$ for any $i \neq j$, is α-balanced for $(\mathcal{M}, \mathcal{D}_A^{(j)})$.

However, given instance $(\mathcal{M}, \mathcal{D}_A^{(j)})$ for an unknown j chosen uniformly at random from $[n]$, with only sample access to $\mathcal{D}_A^{(j)}$, if an algorithm ALG only draws $o(\frac{1}{p_{\min}}) = o(\frac{1}{\delta})$ samples, then with high probability, it cannot distinguish element j from most other elements. As a result, if the input set of active elements A is $[n]$, which is the only case where $j \in A$, ALG will not be able to select element j with constant probability. We prove this formally in Proposition B.1.

2.3 Other useful notions and lemmata

We state two standard concentration inequalities (Lemma A.2) in Section A. Finally, we define a subsampling operator as follows.

---

[4]We assume that matroid $\mathcal{M}$ is given by a membership oracle that answers whether $S \in \mathcal{M}$ for any input set $S \subseteq [n]$, and prior distribution $\mathcal{D}_A$ is given by an oracle that outputs a fresh sample of $\mathcal{D}_A$ upon each query.

Definition 2.8 (subsampling operator $\mathcal{T}_\rho$). *For any $\rho \in [0,1]$, the random operator $\mathcal{T}_\rho$ takes any set $X \subseteq [n]$ as input and outputs a random subset $\mathcal{T}_\rho(X)$ of X such that each element in X appears in $\mathcal{T}_\rho(X)$ independently with probability ρ.*

3   A simple universal OCRS via independent subsampling

In this section, we design a simple universal OCRS with preselected order by subsampling selectable elements independently. Although this subsampling-based OCRS has a slightly weaker universality guarantee than the one in Section 4, its analysis is simpler and of independent interest.

3.1   Main structural lemma

Our universal OCRS is based on the following structural lemma, which says that given a matroid $\mathcal{M}$ and an uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, there exists an element i, such that if we sample a set of active elements A from $\mathcal{D}_A$ and then independently remove each element in A with some constant probability, there is a decent chance that the remaining elements do not span element i conditioned on i being active. This type of lemma was also central to the previous ordered/greedy OCRSs (Chekuri et al., 2014; Feldman et al., 2021) for product distributions.

Lemma 3.1. *For any $\alpha \in [0,1]$ and $\rho \in [0,\alpha]$, given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, there exists $i \in [n]$ such that*

$$\Pr_{A \sim \mathcal{D}_A}[i \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_\rho(A)) \mid i \in A] \geq \alpha - \rho.$$

We note that using subsampling is necessary for the above lemma, in the sense that even if A is sampled from an α-uncontentious distribution $\mathcal{D}_A$ for some strictly positive constant α, it is still possible that every element i is always spanned by $A \setminus \{i\}$ conditioned on $i \in A$.

Example 3.2. *Consider a simple matroid $\mathcal{M} = \{\emptyset, \{1\}, \{2\}\}$ with two elements and a distribution $\mathcal{D}_A$ that is specified by $\Pr[A = \{1,2\}] = \frac{1}{2}$ and $\Pr[A = \emptyset] = \frac{1}{2}$. $\mathcal{D}_A$ is $\frac{1}{2}$-uncontentious because the CRS, that selects an element from 1 and 2 uniformly at random when $A = \{1,2\}$, is $\frac{1}{2}$-balanced. However, we notice that $\Pr_{A \sim \mathcal{D}_A}[i \in \mathrm{span}_{\mathcal{M}}(A \setminus \{i\}) \mid i \in A] = 1$ for any $i \in \{1,2\}$.*

Before proving Lemma 3.1, we establish Lemma 3.3, which is implied by the characterization of uncontentious distributions (Dughmi, 2020, Theorem 2.1). We provide the proof for completeness.

Lemma 3.3. *For any $\alpha \in [0,1]$, for any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, for any weight vector $w \in \mathbb{R}^n_{\geq 0}$, it holds that $\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},w}(A)] \geq \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in A} w_i]$, which in particular, implies that $\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A)] \geq \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[|A|]$.*

Proof. Since $\mathcal{D}_A$ is α-uncontentious for $\mathcal{M}$, there exists an α-balanced CRS $\mathcal{D}_\phi$ for $(\mathcal{M}, \mathcal{D}_A)$. For any weight vector $w \in \mathbb{R}^n_{\geq 0}$, we derive that

$$
\begin{aligned}
\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},w}(A)] &\geq \mathbb{E}_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}[r_{\mathcal{M},w}(\phi(A))] &&\text{(By } \phi(A) \subseteq A \text{ and Lemma A.1-ii.)} \\
&= \mathbb{E}_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}\left[\sum\nolimits_{i \in \phi(A)} w_i\right] &&\text{(By } \phi(A) \in \mathcal{M}) \\
&= \sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}[i \in \phi(A)] \cdot w_i \\
&\geq \sum_{i \in [n]} \alpha \cdot \Pr_{A \sim \mathcal{D}_A}[i \in A] \cdot w_i &&\text{(Since } \mathcal{D}_\phi \text{ is } \alpha\text{-balanced)}
\end{aligned}
$$

$$= \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A} \left[\sum_{i \in A} w_i \right].$$

This implies that $\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A)] \geq \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[|A|]$ by setting w as the all-ones vector. □

Now we proceed to the proof of Lemma 3.1.

Proof of Lemma 3.1. First, we upper bound $\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A)]$ using basic properties of matroids:

$$\begin{aligned}
\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A)] &= \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A \cup \mathcal{T}_\rho(A))] \\
&= \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A \mid \mathcal{T}_\rho(A)) + r_{\mathcal{M}}(\mathcal{T}_\rho(A))] \\
&= \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A \mid \mathcal{T}_\rho(A))] + \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(\mathcal{T}_\rho(A))] \\
&\leq \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in A} r_{\mathcal{M}}(\{i\} \mid \mathcal{T}_\rho(A))] + \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(\mathcal{T}_\rho(A))] && \text{(By Lemma A.1-iii.)} \\
&\leq \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in A} r_{\mathcal{M}}(\{i\} \mid \mathcal{T}_\rho(A))] + \mathbb{E}_{A \sim \mathcal{D}_A}[|\mathcal{T}_\rho(A)|] && \text{(By Lemma A.1-i.)} \\
&= \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in A} r_{\mathcal{M}}(\{i\} \mid \mathcal{T}_\rho(A))] + \rho \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[|A|] && \text{(By definition of } \mathcal{T}_\rho\text{).} \quad (1)
\end{aligned}$$

Since $r_{\mathcal{M}}(\{i\} \mid \mathcal{T}_\rho(A))$ equals 1 if $i \notin \text{span}_{\mathcal{M}}(\mathcal{T}_\rho(A))$ and 0 otherwise, it follows that

$$\begin{aligned}
\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in A} r_{\mathcal{M}}(\{i\} \mid \mathcal{T}_\rho(A))] &= \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in [n]} \mathbb{1}(i \in A, \, i \notin \text{span}_{\mathcal{M}}(\mathcal{T}_\rho(A)))] \\
&= \sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A, \, i \notin \text{span}_{\mathcal{M}}(\mathcal{T}_\rho(A))]. \quad (2)
\end{aligned}$$

Combining Ineq. (1) and Ineq. (2), we have that

$$\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M}}(A)] \leq \sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A, \, i \notin \text{span}_{\mathcal{M}}(\mathcal{T}_\rho(A))] + \rho \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[|A|].$$

By Lemma 3.3, this implies that

$$\sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A, \, i \notin \text{span}_{\mathcal{M}}(\mathcal{T}_\rho(A))] \geq (\alpha - \rho) \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[|A|] = (\alpha - \rho) \cdot \sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A].$$

Therefore, there exists $i \in [n]$ such that

$$\Pr_{A \sim \mathcal{D}_A}[i \notin \text{span}(\mathcal{T}_\rho(A)) \mid i \in A] = \frac{\Pr_{A \sim \mathcal{D}_A}[i \notin \text{span}_{\mathcal{M}}(\mathcal{T}_\rho(A)), \, i \in A]}{\Pr_{A \sim \mathcal{D}_A}[i \in A]} \geq \alpha - \rho.$$

□

3.2 Constructing our universal OCRS with independent subsampling

Our universal OCRS with independent subsampling (Algorithm 1) first samples a set $T = \mathcal{T}_{\frac{\alpha}{2}}([n])$ and applies Lemma 3.1 iteratively to determine an order of the elements, and then following this order, it greedily selects each element that is selectable and belongs to T. In Theorem 3.4, we show that Algorithm 1 is a universal OCRS, and in Section D, we prove that the universality guarantee in Theorem 3.4 is essentially tight for Algorithm 1.

Algorithm 1: UNIVERSAL-OCRS-WITH-INDEPENDENT-SUBSAMPLING

> **Input** : Matroid $\mathcal{M} \subseteq 2^{[n]}$, $\alpha \geq 0$, α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, and $A \sim \mathcal{D}_A$
> **Output:** $X_n \subseteq A$ such that $X_n \in \mathcal{M}$

1 $S_n \leftarrow [n]$, $X_0 \leftarrow \emptyset$, and $T \leftarrow \mathcal{T}_{\frac{\alpha}{2}}([n])$;

2 for $i = n, \ldots, 1$ // notice the reversed order

3 **do**

4 $\quad$ Find an element $\pi(i) \in S_i$ such that $\Pr_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \geq \frac{\alpha}{2}$;

5 $\quad$ $S_{i-1} \leftarrow S_i \setminus \{\pi(i)\}$;

6 end

7 for $i = 1, \ldots, n$ **do**

8 $\quad$ $X_i \leftarrow X_{i-1}$;

9 $\quad$ **if** $\pi(i) \in A \cap T$ **and** $X_{i-1} \cup \{\pi(i)\} \in \mathcal{M}$ **then**

10 $\quad\quad$ $X_i \leftarrow X_{i-1} \cup \{\pi(i)\}$;

11 $\quad$ **end**

12 end

13 return X_n;

Theorem 3.4. *Algorithm 1 is an $(\alpha, \frac{\alpha^2}{4})$-universal OCRS for any $\alpha \in [0, 1]$.*

Proof. For any $\alpha \in [0, 1]$, given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, we can iteratively apply Lemma 3.1 to determine an order $\pi : [n] \to [n]$. Specifically, for each $i \in [n]$, suppose that elements $\pi(i+1), \ldots, \pi(n)$ have already been determined, and let $S_i = [n] \setminus \{\pi(i+1), \ldots, \pi(n)\}$. By Lemma 2.6, since $\mathcal{D}_A$ is α-uncontentious for $\mathcal{M}$, the marginal distribution $\mathcal{D}_A^{S_i}$ is α-uncontentious for the restriction $\mathcal{M}_{S_i}$. We apply Lemma 3.1 to matroid $\mathcal{M}_{S_i}$ and prior distribution $\mathcal{D}_A^{S_i}$ by setting ρ as $\frac{\alpha}{2}$, which implies that there exists an element $\pi(i) \in S_i$ such that $\Pr_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \mathrm{span}_{\mathcal{M}_{S_i}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \geq \frac{\alpha}{2}$. Moreover, by Lemma A.1-vi., $\pi(i) \in \mathrm{span}_{\mathcal{M}_{S_i}}(\mathcal{T}_{\frac{\alpha}{2}}(A'))$ is equivalent to $\pi(i) \in \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A'))$, and thus, we have that

$$\Pr_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \geq \frac{\alpha}{2}. \tag{3}$$

By repeating the argument above (from $i = n$ to $i = 1$), we can determine an order π such that Ineq. (3) holds for all $i \in [n]$ and $S_i = \{\pi(1), \ldots, \pi(i)\}$ (we also let $S_0 = \emptyset$ for completeness), which is exactly what the first for loop of Algorithm 1 does.

Now given a set of active elements $A \sim \mathcal{D}_A$, consider the selection procedure in the second for loop of Algorithm 1. For each $i \in [n]$, element $\pi(i)$ is included in X_i only if $\pi(i) \in A \cap T$. Therefore, for all $i \in [n]$, the solution set X_{i-1} is a subset of $A \cap T \cap S_{i-1}$, and hence, we have that

$$\begin{aligned}
&\Pr_{A \sim \mathcal{D}_A}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(X_{i-1}) \mid \pi(i) \in A] \\
&\geq \Pr_{A \sim \mathcal{D}_A}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(A \cap T \cap S_{i-1}) \mid \pi(i) \in A] &&\text{(By Lemma A.1-iv.)} \\
&= \Pr_{A \sim \mathcal{D}_A}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A \cap S_{i-1})) \mid \pi(i) \in A] &&\text{(Since } T = \mathcal{T}_{\frac{\alpha}{2}}([n])\text{)} \\
&\geq \Pr_{A \sim \mathcal{D}_A}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A \cap S_i)) \mid \pi(i) \in A] &&\text{(Since } S_{i-1} \subseteq S_i\text{)} \\
&= \Pr_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] &&\text{(By definition of } \mathcal{D}_A^{S_i}\text{).} \tag{4}
\end{aligned}$$

Moreover, because element $\pi(i)$ is selected by Algorithm 1 if $\pi(i) \in A \cap T$ and $\pi(i) \notin \mathrm{span}_{\mathcal{M}}(X_{i-1})$, the probability that $\pi(i)$ is selected conditioned on it being active is

$$\Pr_{A \sim \mathcal{D}_A}[\pi(i) \in T, \ \pi(i) \notin \mathrm{span}_{\mathcal{M}}(X_{i-1}) \mid \pi(i) \in A]$$

8

$$= \Pr_{A \sim \mathcal{D}_A}[\pi(i) \notin \text{span}_{\mathcal{M}}(X_{i-1}) \mid \pi(i) \in A] \times \Pr[\pi(i) \in T] \qquad \text{(Since } T \text{ is independent of } A)$$

$$= \Pr_{A \sim \mathcal{D}_A}[\pi(i) \notin \text{span}_{\mathcal{M}}(X_{i-1}) \mid \pi(i) \in A] \times \frac{\alpha}{2} \qquad \text{(Since } T = \mathcal{T}_{\frac{\alpha}{2}}([n]))$$

$$\geq \Pr_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \text{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \times \frac{\alpha}{2} \qquad \text{(By Ineq. (4))} \qquad (5)$$

$$\geq \frac{\alpha}{2} \times \frac{\alpha}{2} = \frac{\alpha^2}{4} \qquad \text{(By Ineq. (3)).}$$

It follows that Algorithm 1 is $\frac{\alpha^2}{4}$-balanced for $(\mathcal{M}, \mathcal{D}_A)$, which finishes the proof. $\qquad \square$

We note that Algorithm 1 does not specify how to find element $\pi(i)$ at Line 4. In Section C, we implement this step by estimating probabilities $\Pr_{A' \sim \mathcal{D}_A^{S_i}}[j \in \text{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid j \in A']$ for all $i \in [n]$ and $j \in S_i$ using Monte-Carlo sampling, which results in the following corollary.

Corollary 3.5. *For any* $\alpha, \varepsilon \in (0, 1]$, *there exists an* $(\alpha, \frac{(1-\varepsilon)\alpha^2}{4})$-*universal OCRS with preselected order, which given input matroid* $\mathcal{M} \subseteq 2^{[n]}$ *and* α-*uncontentious prior distribution* $\mathcal{D}_A$ *for* $\mathcal{M}$, *runs in* $O\left(\frac{n \log(n/\varepsilon)}{\alpha^2 \varepsilon^2 p_{\min}} \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}} \cdot n)\right)$ *time, where* $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$, *and* $t_{\mathcal{D}_A}, t_{\mathcal{M}}$ *are the time it takes to generate a sample from* $\mathcal{D}_A$ *and to check whether a set of elements belongs to* $\mathcal{M}$ *respectively.*

Furthermore, as we mentioned in the introduction, Algorithm 1 (in fact, all of our universal OCRSs) generalizes ordered OCRSs for product distributions (Chekuri et al., 2014; Gupta and Nagarajan, 2013). For product distributions, it is known that ordered OCRSs can be strengthened to obtain *greedy* OCRSs, which work even for the worst-case arrival model (Feldman et al., 2021, Theorem 2.1). One might wonder whether we can also apply our techniques to make those greedy OCRSs universal. Unfortunately, the answer is no, because this would yield a universal OCRS for the worst-case arrival model, which does not exist even for 1-uniform matroids (Dughmi, 2020, Theorem 5.7). However, this does not preclude the possibility that there might be meaningful relaxations of greedy OCRSs which can be made universal.

4 An improved universal OCRS via correlated subsampling

In this section, we design a universal OCRS with preselected order using correlated subsampling, which achieves a slightly stronger universality guarantee than Algorithm 1. We first introduce some notations which we will use in this section. We will use symbols π, σ, τ to represent permutations. Given any permutation $\sigma : [n] \to [n]$ and any element $e \in [n]$, we let $\text{prefix}(\sigma, e)$ denote the set of elements appearing before element e in permutation σ, namely, $\text{prefix}(\sigma, e) := \{\sigma(j) \mid j \in [\sigma^{-1}(e) - 1]\}$. Moreover, for any $k \in [n]$, we let σ^k denote the set of the first k elements in permutation σ, namely, $\sigma^k := \{\sigma(j) \mid j \in [k]\}$, and we let $\sigma^0 := \emptyset$. Furthermore, for any set of elements S, we let $\mathcal{P}(S)$ denote the uniform distribution over all permutations of elements in S.

4.1 Main structural lemma

The main idea of our improved universal OCRS is replacing the independent subsampling operator $\mathcal{T}_\rho$ with a correlated subsampling method[5] that is inspired by Lemma 4.1. Essentially, Lemma 4.1 says that given a matroid $\mathcal{M}$ and an uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, there exists an element i, such that if we sample a set of active elements A from $\mathcal{D}_A$ and a uniformly random permutation

---

[5]The author thanks an anonymous reviewer for suggesting this method and raising valuable questions.

9

σ, and then remove all elements in A except those appearing before element i in permutation σ, there is a decent chance that the remaining elements do not span i conditioned on i being active.

Lemma 4.1. *For any $\alpha \in [0, 1]$, given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, there exists some $i \in [n]$ such that*

$$\Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[i \notin \mathrm{span}(A \cap \mathsf{prefix}(\sigma, i)) \mid i \in A] \geq \alpha.$$

Proof. Since $\mathcal{D}_A$ is α-uncontentious for $\mathcal{M}$, there exists an α-balanced CRS $\mathcal{D}_\phi$ for $(\mathcal{M}, \mathcal{D}_A)$. First, we notice that Lemma 3.3 implies that

$$\mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[r_\mathcal{M}(A)] \geq \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[|A|], \tag{6}$$

because $r_\mathcal{M}(A)$ and $|A|$ are independent of σ. We decompose $\mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[r_\mathcal{M}(A)]$ as follows,

$$\mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[r_\mathcal{M}(A)]$$
$$= \mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}\Big[\sum_{k \in [n]} r_\mathcal{M}(A \cap \sigma^k) - r_\mathcal{M}(A \cap \sigma^{k-1})\Big]$$
$$\quad \text{(Telescoping sum)}$$
$$= \mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}\Big[\sum_{k \in [n]} r_\mathcal{M}(A \cap (\mathsf{prefix}(\sigma, \sigma(k)) \cup \{\sigma(k)\})) - r_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, \sigma(k)))\Big]$$
$$\quad \text{(Since } \mathsf{prefix}(\sigma, \sigma(k)) = \sigma^{k-1} \text{ and } \sigma^k = \sigma^{k-1} \cup \{\sigma(k)\})$$
$$= \mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}\Big[\sum_{i \in [n]} r_\mathcal{M}(A \cap (\mathsf{prefix}(\sigma, i) \cup \{i\})) - r_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))\Big]$$
$$\quad \text{(Since } \sigma \text{ is a permutation)}$$
$$= \sum_{i \in [n]} \mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[r_\mathcal{M}(A \cap (\mathsf{prefix}(\sigma, i) \cup \{i\})) - r_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))]. \tag{7}$$

Note that the term $r_\mathcal{M}(A \cap (\mathsf{prefix}(\sigma, i) \cup \{i\})) - r_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))$ equals 1 if $i \in A$ and $i \notin \mathrm{span}_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))$, and equals 0 otherwise. It follows that

$$\mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[r_\mathcal{M}(A \cap (\mathsf{prefix}(\sigma, i) \cup \{i\})) - r_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))]$$
$$= \Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[i \in A, i \notin \mathrm{span}_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))].$$

Combining this with Eq. (7), we get

$$\mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[r_\mathcal{M}(A)] = \sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[i \in A, i \notin \mathrm{span}_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))]. \tag{8}$$

Furthermore, we can decompose $\mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[|A|]$ as follows,

$$\mathbb{E}_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[|A|] = \sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[i \in A]. \tag{9}$$

Putting Ineq. (6), Eq. (8) and Eq. (9) together, we get

$$\sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[i \in A, i \notin \mathrm{span}_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))] \geq \alpha \cdot \sum_{i \in [n]} \Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[i \in A],$$

10

which implies that there exists $i \in [n]$ such that

$$\Pr_{\substack{A \sim \mathcal{D}_A, \\ \sigma \sim \mathcal{P}([n])}}[i \notin \mathsf{span}_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i)) \mid i \in A] = \frac{\Pr_{\substack{A \sim \mathcal{D}_A, \\ \sigma \sim \mathcal{P}([n])}}[i \in A, \ i \notin \mathsf{span}_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, i))]}{\Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}([n])}[i \in A]} \geq \alpha.$$

$\square$

We can iteratively apply Lemma 4.1 to determine an order of elements, which will be the preselected order of our improved universal OCRS. We formally describe this procedure in Subroutine 2 and state its guarantee in Corollary 4.2.

Subroutine 2: ORDER-PRESELECTING

Input : Matroid $\mathcal{M} \subseteq 2^{[n]}$, $\alpha \geq 0$, α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$
Output: Permutation $\pi : [n] \to [n]$

1   $S_n \leftarrow [n]$;
2   **for** $i = n, \ldots, 1$ // notice the reversed order
3   **do**
4   Find an element $\pi(i) \in S_i$ such that
      $\Pr_{A' \sim \mathcal{D}_A^{S_i}, \sigma \sim \mathcal{P}(S_i)}[\pi(i) \notin \mathsf{span}_\mathcal{M}(A' \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A'] \geq \alpha$;
5      $S_{i-1} \leftarrow S_i \setminus \{\pi(i)\}$
6   **end**
7   **return** π;

Corollary 4.2. *For any $\alpha \in [0, 1]$, given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, Subroutine 2 outputs a permutation $\pi : [n] \to [n]$ such that for all $i \in [n]$,*

$$\Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \mathsf{span}_\mathcal{M}(A \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A] \geq \alpha. \tag{10}$$

Proof. For any $\alpha \in [0, 1]$, given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, we can iteratively apply Lemma 4.1 to determine an order $\pi : [n] \to [n]$. Specifically, for each $i \in [n]$, suppose that elements $\pi(i + 1), \ldots, \pi(n)$ have already been determined, and let $S_i = [n] \setminus \{\pi(i + 1), \ldots, \pi(n)\}$. By Lemma 2.6, since $\mathcal{D}_A$ is α-uncontentious for $\mathcal{M}$, the marginal distribution $\mathcal{D}_A^{S_i}$ is α-uncontentious for the restriction $\mathcal{M}_{S_i}$. We apply Lemma 4.1 to matroid $\mathcal{M}_{S_i}$ and prior distribution $\mathcal{D}_A^{S_i}$, which implies that there exists an element $\pi(i) \in S_i$ such that $\Pr_{A' \sim \mathcal{D}_A^{S_i}, \sigma \sim \mathcal{P}(S_i)}[\pi(i) \notin \mathsf{span}_{\mathcal{M}_{S_i}}(A' \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A'] \geq \alpha$. Moreover, by Lemma A.1-vi., $\pi(i) \notin \mathsf{span}_{\mathcal{M}_{S_i}}(A' \cap \mathsf{prefix}(\sigma, \pi(i)))$ is equivalent to $\pi(i) \notin \mathsf{span}_\mathcal{M}(A' \cap \mathsf{prefix}(\sigma, \pi(i)))$, and hence, we have that $\Pr_{A' \sim \mathcal{D}_A^{S_i}, \sigma \sim \mathcal{P}(S_i)}[\pi(i) \notin \mathsf{span}_\mathcal{M}(A' \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A'] \geq \alpha$.

By repeating the argument above (from $i = n$ to $i = 1$), we can determine an order π such that for all $i \in [n]$, $\Pr_{A' \sim \mathcal{D}_A^{S_i}, \sigma \sim \mathcal{P}(S_i)}[\pi(i) \notin \mathsf{span}_\mathcal{M}(A' \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A'] \geq \alpha$, which is exactly what Subroutine 2 does. This implies Ineq. (10) because we have that

$$\Pr_{A' \sim \mathcal{D}_A^{S_i}, \sigma \sim \mathcal{P}(S_i)}[\pi(i) \notin \mathsf{span}_\mathcal{M}(A' \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A']$$

$$= \Pr_{A' \sim \mathcal{D}_A^{\pi^i}, \sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \mathsf{span}_\mathcal{M}(A' \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A'] \qquad \text{(Since } S_i = \pi^i\text{)}$$

$$= \Pr_{A \sim \mathcal{D}_A, \sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \mathsf{span}_\mathcal{M}(A \cap \pi^i \cap \mathsf{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A \cap \pi^i] \quad \text{(By definition of } \mathcal{D}_A^{\pi^i}\text{)}$$

$$= \Pr_{A \sim \mathcal{D}_A, \, \sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A],$$

where the last equality follows from the facts that $\text{prefix}(\sigma, \pi(i)) \subseteq \pi^i$ and that $\pi(i) \in \pi^i$. $\qquad\square$

4.2 Correlated subsampling

In order to make use of Corollary 4.2, given a permutation π generated by Subroutine 2, we need to sample a set $T \subseteq [n]$ such that for each $i \in [n]$, the subset $\pi^{i-1} \cap T$ follows the same distribution as the random set $\text{prefix}(\sigma, \pi(i))$, where $\sigma \sim \mathcal{P}(\pi^i)$. In Lemma 4.3, we show that this can be achieved by first sampling a permutation $\sigma' \sim \mathcal{P}([n+1])$ and then setting T as $\text{prefix}(\sigma', n+1)$.

Lemma 4.3. *Suppose that we are given a permutation $\pi : [n] \to [n]$, and we sample a permutation $\sigma' \sim \mathcal{P}([n+1])$ and let $T = \text{prefix}(\sigma', n+1)$. Then, for all $i \in [n]$, the subset $T_{i-1} := \pi^{i-1} \cap T$ follows the same distribution as the random set $\text{prefix}(\sigma, \pi(i))$, where $\sigma \sim \mathcal{P}(\pi^i)$. Moreover, for all $i \in [n]$ and $S \subseteq \pi^{i-1}$, we have that*

$$\Pr[\pi(i) \in T \mid T_{i-1} = S] = \frac{|S|+1}{i+1}. \tag{11}$$

Proof. Given any permutation $\pi : [n] \to [n]$, for any $i \in [n]$, we show that the outputs of the following three random processes have the same distribution:

(1) Sample $\sigma \sim \mathcal{P}(\pi^i)$ and output $\text{prefix}(\sigma, \pi(i))$.

(2) Sample $\sigma'' \sim \mathcal{P}(\pi^{i-1} \cup \{n+1\})$ and output $\text{prefix}(\sigma'', n+1)$.

(3) Sample $\sigma' \sim \mathcal{P}([n+1])$ and output $\text{prefix}(\sigma', n+1) \cap (\pi^{i-1} \cup \{n+1\})$.

Specifically, the second random process is equivalent to the first, because it merely relabels $\pi(i)$ with $n+1$. The third random process is equivalent to the second, because sampling a uniformly random permutation of $\pi^{i-1} \cup \{n+1\}$ is the same as first sampling a uniformly random permutation of $[n+1]$ and then keeping only the elements of $\pi^{i-1} \cup \{n+1\}$. Thus, the outputs of all three processes follow the same distribution.

Moreover, notice that the output of the third process is equal to $\text{prefix}(\sigma', n+1) \cap \pi^{i-1}$ because $n+1 \notin \text{prefix}(\sigma', n+1)$. We let $T = \text{prefix}(\sigma', n+1)$ and $T_{i-1} = T \cap \pi^{i-1}$ as in the lemma statement. Then, the subset T_{i-1}, which is equal to the output of the third process, follows the same distribution as the output of the first process, which is $\text{prefix}(\sigma, \pi(i))$.

To prove Eq. (11), we let $T' = \text{prefix}(\sigma', n+1) \cap (\pi^i \cup \{n+1\})$. Analogous to the equivalence between the second and the third random processes above, we observe that T' follows the same distribution as the random set $\text{prefix}(\tau, n+1)$, where $\tau \sim \mathcal{P}(\pi^i \cup \{n+1\})$. Furthermore, we note that $\pi(i) \in T$ iff $\pi(i) \in T'$, and that $\pi^{i-1} \cap T = \pi^{i-1} \cap T'$. Hence, we have that for any $S \subseteq \pi^{i-1}$,

$$\Pr[\pi(i) \in T \mid \pi^{i-1} \cap T = S] = \Pr[\pi(i) \in T' \mid \pi^{i-1} \cap T' = S]$$
$$= \Pr_{\tau \sim \mathcal{P}(\pi^i \cup \{n+1\})}[\pi(i) \in \text{prefix}(\tau, n+1) \mid \pi^{i-1} \cap \text{prefix}(\tau, n+1) = S]. \tag{12}$$

We observe that sampling $\tau \sim \mathcal{P}(\pi^i \cup \{n+1\})$ is equivalent to the following sampling process: First, sample a permutation $\tau \sim \mathcal{P}(\pi^{i-1} \cup \{n+1\})$. Then, sample $j \in [i+1]$ uniformly at random. If $j < i+1$, insert element $\pi(i)$ in front of the j-th element in permutation τ. If $j = i+1$, insert element $\pi(i)$ after the last element in permutation τ.

Notice that in the second step of this sampling process, there are $|\pi^{i-1} \cap \mathsf{prefix}(\tau, n+1)| + 1$ positions before element $n+1$ and $i+1$ positions in total, where element $\pi(i)$ can possibly be inserted. It follows that for any $S \subseteq \pi^{i-1}$,

$$\Pr_{\tau \sim \mathcal{P}(\pi^i \cup \{n+1\})}[\pi(i) \in \mathsf{prefix}(\tau, n+1) \mid \pi^{i-1} \cap \mathsf{prefix}(\tau, n+1) = S] = \frac{|S|+1}{i+1},$$

and hence, by Eq. (12), we have that

$$\Pr[\pi(i) \in T \mid \pi^{i-1} \cap T = S] = \frac{|S|+1}{i+1},$$

which implies Eq. (11) since $T_{i-1} = \pi^{i-1} \cap T$. □

4.3 Constructing our universal OCRS with correlated subsampling

We are ready to present our improved universal OCRS with correlated subsampling (Algorithm 3). Algorithm 3 first runs Subroutine 2 to preselect an order π, and then, it samples a set T according to Lemma 4.3. Finally, it iterates through all elements following order π, and for each $i \in [n]$, it selects element $\pi(i)$ if $\pi(i)$ is selectable and belongs to T.

Algorithm 3: UNIVERSAL-OCRS-WITH-CORRELATED-SUBSAMPLING

Input : Matroid $\mathcal{M} \subseteq 2^{[n]}$, $\alpha \geq 0$, α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, and $A \sim \mathcal{D}_A$

Output: $X_n \subseteq A$ such that $X_n \in \mathcal{M}$

1 Run Subroutine 2 on input $(\mathcal{M}, \alpha, \mathcal{D}_A)$, which outputs a permutation $\pi : [n] \to [n]$;

2 Sample a permutation $\sigma' \sim \mathcal{P}([n+1])$;

3 $T \leftarrow \mathsf{prefix}(\sigma', n+1)$ and $X_0 \leftarrow \emptyset$;

4 for $i = 1, \ldots, n$ **do**

5 $\quad X_i \leftarrow X_{i-1}$;

6 $\quad$ **if** $\pi(i) \in A \cap T$ **and** $X_{i-1} \cup \{\pi(i)\} \in \mathcal{M}$ **then**

7 $\quad\quad \mid X_i \leftarrow X_{i-1} \cup \{\pi(i)\}$;

8 $\quad$ **end**

9 end

10 return X_n;

In Theorem 4.4, we show that Algorithm 3 is a universal OCRS, and in Section D, we prove that the universality guarantee in Theorem 4.4 is essentially tight for Algorithm 3.

Theorem 4.4. *Algorithm 3 is an $(\alpha, \frac{\alpha^2}{2})$-universal OCRS for any $\alpha \in [0,1]$.*

Before we prove Theorem 4.4, we note that unlike Algorithm 1, which is based on independent subsampling, Algorithm 3 employs a correlated subsampling method. Therefore, in Algorithm 3, the event that element $\pi(i)$ is subsampled (i.e., $\pi(i) \in T$) is correlated with the event that element $\pi(i)$ is spanned by the partial solution set X_{i-1} (which is a subset of T). Analyzing this correlation will be the main technicality of the proof. In the following, we state a useful inequality.

Fact 4.5. *For any integer $i \geq 1$, for any real numbers $\alpha, x_0, \ldots, x_{i-1} \in [0,1]$, if $\sum_{k=0}^{i-1} \frac{1}{i} \cdot x_k \geq \alpha$, then $\sum_{k=0}^{i-1} \frac{k+1}{i(i+1)} \cdot x_k \geq \frac{\alpha^2}{2}$.*

We defer the proof of Fact 4.5 to Section B and proceed to the proof of Theorem 4.4.

Proof of Theorem 4.4. For any $\alpha \in [0,1]$, given any matroid $\mathcal{M}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, we want to show that for all $i \in [n]$, Algorithm 3 selects element $\pi(i)$ with probability at least $\frac{\alpha^2}{2}$ conditioned on $\pi(i)$ being active. We denote $T_i := \pi^i \cap T$ for all $i \in \{0, 1, \ldots, n\}$. For readability, we will omit the random sets A, T and T_i's from the subscripts of Pr in the proof.

First, notice that by Line 6, for all $i \in [n]$, when element $\pi(i)$ is active, it is selected if $\pi(i) \in T_i$ and $\pi(i) \notin \text{span}_{\mathcal{M}}(X_{i-1})$. We observe that $X_{i-1} \subseteq A \cap T_{i-1}$, and thus, by Lemma A.1-iv., $\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1})$ implies $\pi(i) \notin \text{span}_{\mathcal{M}}(X_{i-1})$. Hence, for any $i \in [n]$, to prove that element $\pi(i)$ is selected with probability at least $\frac{\alpha^2}{2}$ conditioned on it being active, it suffices to show that

$$\Pr[\pi(i) \in T_i,\, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A] \geq \frac{\alpha^2}{2}. \tag{13}$$

To this end, for any $i \in [n]$, we first decompose $\Pr[\pi(i) \in T_i,\, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A]$ as

$$\Pr[\pi(i) \in T_i,\, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A]$$

$$= \sum_{S \subseteq \pi^{i-1}} \Pr[T_{i-1} = S,\, \pi(i) \in T_i,\, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S) \mid \pi(i) \in A]$$

$$\text{(Since } T_{i-1} \subseteq \pi^{i-1})$$

$$= \sum_{S \subseteq \pi^{i-1}} \Pr[T_{i-1} = S \mid \pi(i) \in A] \times \Pr[\pi(i) \in T_i,\, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S) \mid \pi(i) \in A,\, T_{i-1} = S]$$

$$= \sum_{S \subseteq \pi^{i-1}} \Pr[T_{i-1} = S] \times \Pr[\pi(i) \in T_i,\, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S) \mid \pi(i) \in A,\, T_{i-1} = S]$$

$$\text{(Since } T_{i-1} \text{ is independent of } A)$$

$$= \sum_{S \subseteq \pi^{i-1}} \Pr[T_{i-1} = S] \times \Pr[\pi(i) \in T_i \mid \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S),\, \pi(i) \in A,\, T_{i-1} = S]$$

$$\times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A,\, T_{i-1} = S]$$

$$= \sum_{S \subseteq \pi^{i-1}} \Pr[T_{i-1} = S] \times \Pr[\pi(i) \in T_i \mid T_{i-1} = S]$$

$$\times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A,\, T_{i-1} = S]$$

$$\text{(Since } T_i \text{ is independent of } A)$$

$$= \sum_{S \subseteq \pi^{i-1}} \Pr[T_{i-1} = S] \times \frac{|S| + 1}{i + 1} \times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S) \mid \pi(i) \in A,\, T_{i-1} = S]$$

$$\text{(By Eq. (11) in Lemma 4.3).}$$

Then, we group the subsets $S \subseteq \pi^{i-1}$ according to their sizes, and we have that

$$\Pr[\pi(i) \in T_i,\, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A]$$

$$= \sum_{k=0}^{i-1} \sum_{\substack{S \subseteq \pi^{i-1} \\ \text{s.t. } |S| = k}} \Pr[T_{i-1} = S] \times \frac{k+1}{i+1} \times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S) \mid \pi(i) \in A,\, T_{i-1} = S]$$

$$= \sum_{k=0}^{i-1} \sum_{\substack{S \subseteq \pi^{i-1} \\ \text{s.t. } |S| = k}} \frac{k+1}{i+1} \times \Pr[T_{i-1} = S \mid \pi(i) \in A] \times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S) \mid \pi(i) \in A,\, T_{i-1} = S]$$

$$\text{(Since } T_{i-1} \text{ is independent of } A)$$

14

$$= \sum_{k=0}^{i-1} \sum_{\substack{S \subseteq \pi^{i-1} \\ \text{s.t. } |S|=k}} \frac{k+1}{i+1} \times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap S), T_{i-1} = S \mid \pi(i) \in A]$$

$$= \sum_{k=0}^{i-1} \frac{k+1}{i+1} \times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}), |T_{i-1}| = k \mid \pi(i) \in A]$$

$$= \sum_{k=0}^{i-1} \frac{k+1}{i+1} \times \Pr[|T_{i-1}| = k \mid \pi(i) \in A] \times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A, |T_{i-1}| = k]$$

$$= \sum_{k=0}^{i-1} \frac{k+1}{i+1} \times \Pr[|T_{i-1}| = k] \times \Pr[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A, |T_{i-1}| = k]$$

(Since T_{i-1} is independent of A)

$$= \sum_{k=0}^{i-1} \frac{k+1}{i+1} \times \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[|\text{prefix}(\sigma, \pi(i))| = k]$$
$$\times \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A, |\text{prefix}(\sigma, \pi(i))| = k]$$

(By Lemma 4.3, T_{i-1} follows the same distribution as $\text{prefix}(\sigma, \pi(i))$, where $\sigma \sim \mathcal{P}(\pi^i)$)

$$= \sum_{k=0}^{i-1} \frac{k+1}{i+1} \times \frac{1}{i} \times \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A, |\text{prefix}(\sigma, \pi(i))| = k], \quad (14)$$

where the last equality is because the number of elements that precede element $\pi(i)$ in a uniformly random permutation of π^i follows the uniform distribution over $\{0, \ldots, i-1\}$. For simplicity, we denote $P_{i,k} := \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A, |\text{prefix}(\sigma, \pi(i))| = k]$ for all $i \in [n]$ and $k \in \{0, \ldots, i-1\}$, and then Eq. (14) simplifies to

$$\Pr[\pi(i) \in T_i, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap T_{i-1}) \mid \pi(i) \in A] = \sum_{k=0}^{i-1} \frac{k+1}{i(i+1)} \cdot P_{i,k}. \quad (15)$$

Next, we decompose $\Pr_{\sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A]$ as follows,

$$\Pr_{\sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A]$$

$$= \sum_{k=0}^{i-1} \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[|\text{prefix}(\sigma, \pi(i))| = k, \pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A]$$

(The number of elements preceding element $\pi(i)$ in σ is at most $|\pi^i| - 1 = i - 1$)

$$= \sum_{k=0}^{i-1} \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[|\text{prefix}(\sigma, \pi(i))| = k \mid \pi(i) \in A]$$
$$\times \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A, |\text{prefix}(\sigma, \pi(i))| = k]$$

$$= \sum_{k=0}^{i-1} \Pr_{\sigma \sim \mathcal{P}(\pi^i)}[|\text{prefix}(\sigma, \pi(i))| = k] \times P_{i,k}$$

(By definition of $P_{i,k}$ and the fact that $\text{prefix}(\sigma, \pi(i))$ is independent of A)

15

$$= \sum_{k=0}^{i-1} \frac{1}{i} \cdot P_{i,k}. \tag{16}$$

Because Corollary 4.2 guarantees that $\Pr_{\sigma \sim \mathcal{P}(\pi^i)}[\pi(i) \notin \text{span}_{\mathcal{M}}(A \cap \text{prefix}(\sigma, \pi(i))) \mid \pi(i) \in A] \geq \alpha$, by Eq. (16), we have that $\sum_{k=0}^{i-1} \frac{1}{i} \cdot P_{i,k} \geq \alpha$. Finally, it follows by Fact 4.5 that $\sum_{k=0}^{i-1} \frac{k+1}{i(i+1)} \cdot P_{i,k} \geq \frac{\alpha^2}{2}$, which implies Ineq. (13) by Eq. (15). $\qquad \square$

5 An optimal universal OCRS via linear programming

In this section, we present an LP that computes a universal OCRS with nearly optimal universality guarantee in the preselected-order model, and we solve this LP efficiently using the ellipsoid method. In fact, this approach was originally developed by Chekuri et al. (2014) to compute optimal offline CRSs for product distributions. We show that it applies naturally to ordered OCRSs for (arbitrarily correlated) uncontentious distributions.

Algorithm 4: Deterministic ordered OCRS ϕ_π

    Input : Matroid $\mathcal{M} \subseteq 2^{[n]}$, prior distribution $\mathcal{D}_A$, and $A \sim \mathcal{D}_A$
    Output: $X \subseteq A$ such that $X \in \mathcal{M}$
1 $X \leftarrow \emptyset$;
2 **for** $i = 1, \ldots, n$ **do**
3   **if** $\pi(i) \in A$ **and** $X \cup \{\pi(i)\} \in \mathcal{M}$ **then**
4 $X \leftarrow X \cup \{\pi(i)\}$;
5   **end**
6 **end**
7 **return** X;

We start by introducing some notations which we will use in this section. We let $\mathcal{S}_n$ denote the set of all permutations of $[n]$. For each permutation $\pi \in \mathcal{S}_n$, we let ϕ_π denote the deterministic ordered OCRS with preselected order π (Algorithm 4). Moreover, we define a permutation π_w for each weight vector $w \in \mathbb{R}_{\geq 0}^n$ as follow:

Definition 5.1. *Given any weight vector $w \in \mathbb{R}_{\geq 0}^n$, we assign weight w_i to each element $i \in [n]$, and we permute elements of $[n]$ in the decreasing order of their weights (with arbitrary tie-breaking for equal weights), and then we let π_w denote this permutation.*

In the following lemma, we establish a key property of OCRS ϕ_π.

Lemma 5.2. *For any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any prior distribution $\mathcal{D}_A \in \Delta(2^{[n]})$, for any weight vector $w \in \mathbb{R}_{\geq 0}^n$ and any permutation $\pi \in \mathcal{S}_n$, we have that*

$$\mathbb{E}_{A \sim \mathcal{D}_A} \left[\sum_{i \in \phi_\pi(A)} w_i \right] \leq \mathbb{E}_{A \sim \mathcal{D}_A} \left[\sum_{i \in \phi_{\pi_w}(A)} w_i \right] = \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},w}(A)].$$

Proof. Given any weight vector $w \in \mathbb{R}_{\geq 0}^n$, we assign weight w_i to each element $i \in [n]$. We first prove that $\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_{\pi_w}(A)} w_i] = \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},w}(A)]$. We observe that given a set of active elements $A \subseteq [n]$ as input, OCRS ϕ_{π_w} visits elements of A in the decreasing order of their weights and greedily selects each element that can be selected without violating the matroid constraint. It is well-known that this greedy procedure selects a maximum-weight subset of A that satisfies the

matroid constraint (see e.g., Welsh (2010, Chapter 19)). Therefore, for any $A \subseteq [n]$, we have that $\sum_{i \in \phi_{\pi_w}(A)} w_i = r_{\mathcal{M},w}(A)$. It follows that $\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_{\pi_w}(A)} w_i] = \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},w}(A)]$.

Next, we show that for any permutation $\pi \in \mathcal{S}_n$, $\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_{\pi}(A)} w_i] \leq \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},w}(A)]$. We observe that given any input set of active elements $A \subseteq [n]$, OCRS ϕ_π guarantees that $\phi_\pi(A) \subseteq A$ and $\phi_\pi(A) \in \mathcal{M}$. Therefore, for any $A \subseteq [n]$, we have that $\sum_{i \in \phi_\pi(A)} w_i \leq r_{\mathcal{M},w}(A)$. It follows that $\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_\pi(A)} w_i] \leq \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},w}(A)]$. $\qquad\square$

5.1 Formulating the LP

Now we formulate an LP that computes an α-balanced OCRS for any matroid $\mathcal{M}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$. The resulting OCRS will be a random mixture of OCRSs ϕ_π for various permutations $\pi \in \mathcal{S}_n$. Specifically, we let $x_i := \Pr_{A \sim \mathcal{D}_A}[i \in A]$ and $q_{i,\pi} := \Pr_{A \sim \mathcal{D}_A}[i \in \phi_\pi(A)]$ for all $i \in [n]$ and $\pi \in \mathcal{S}_n$, and we consider the following LP (LP) and its dual (DP).

$$
\begin{aligned}
\text{(LP)} \quad & \max_{\beta, \lambda_\pi} \beta \\
& \text{s.t.} \sum_{\pi \in \mathcal{S}_n} q_{i,\pi} \lambda_\pi \geq \beta x_i && \forall i \in [n] \\
& \qquad \sum_{\pi \in \mathcal{S}_n} \lambda_\pi = 1 \\
& \qquad \lambda_\pi \geq 0 && \forall \pi \in \mathcal{S}_n.
\end{aligned}
$$

$$
\begin{aligned}
\text{(DP)} \quad & \min_{\gamma, \mu_i} \gamma \\
& \text{s.t.} \sum_{i \in [n]} q_{i,\pi} \mu_i \leq \gamma && \forall \pi \in \mathcal{S}_n \\
& \qquad \sum_{i \in [n]} x_i \mu_i = 1 \\
& \qquad \mu_i \geq 0 && \forall i \in [n]. \qquad (17)
\end{aligned}
$$

We note that every feasible solution (β, λ_π) to (LP) corresponds to a β-balanced OCRS with preselected order for $(\mathcal{M}, \mathcal{D}_A)$. Indeed, because of the last two constraints in (LP), variables λ_π for all $\pi \in \mathcal{S}_n$ together specify a distribution $\mathcal{D}_\pi$ over $\mathcal{S}_n$. We observe that ϕ_π with $\pi \sim \mathcal{D}_\pi$ is a randomized OCRS with preselected order for $(\mathcal{M}, \mathcal{D}_A)$, and the first constraint in (LP) ensures that this OCRS is β-balanced.

Next, we show that if the prior distribution $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, then the optimal value of (DP) is at least α. By LP duality, the optimal value of (LP) is also at least α, and hence, the optimal solution to (LP) corresponds to an α-balanced OCRS with preselected order for $(\mathcal{M}, \mathcal{D}_A)$.

Lemma 5.3. *For any $\alpha \in [0,1]$, if the prior distribution $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, then any vector $\mu \in \mathbb{R}^n$ that satisfies the last two constraints in (DP) in Eq. (17) must also satisfy that $\sum_{i \in [n]} q_{i,\pi_\mu} \mu_i \geq \alpha$, where permutation π_μ is defined in Definition 5.1. This implies that the optimal value of (DP) is at least α.*

Proof. For any vector $\mu \in \mathbb{R}^n$ that satisfies the last two constraints in (DP), by Lemma 5.2, we have that $\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_{\pi_\mu}(A)} \mu_i] = \mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},\mu}(A)]$. Since $\mathcal{D}_A$ is α-uncontentious, by Lemma 3.3, we have that $\mathbb{E}_{A \sim \mathcal{D}_A}[r_{\mathcal{M},\mu}(A)] \geq \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in A} \mu_i]$. It follows that $\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_{\pi_\mu}(A)} \mu_i] \geq \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in A} \mu_i]$, which is equivalent to $\sum_{i \in [n]} q_{i,\pi_\mu} \mu_i \geq \alpha \cdot \sum_{i \in [n]} x_i \mu_i$. Notice that $\sum_{i \in [n]} x_i \mu_i = 1$

by the second constraint in (DP). Therefore, we have that $\sum_{i \in [n]} q_{i,\pi_\mu} \mu_i \geq \alpha$ for any vector μ that satisfies the last two constraints in (DP). This implies that any feasible solution (γ, μ_i) to (DP) must satisfy $\gamma \geq \alpha$ because of the constraint $\sum_{i \in [n]} q_{i,\pi_\mu} \mu_i \leq \gamma$. It follows that the optimal value of (DP) is at least α. $\qquad\square$

5.2 Solving the LP efficiently

We have shown that for any matroid $\mathcal{M}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, we can find an α-balanced OCRS for $(\mathcal{M}, \mathcal{D}_A)$ by solving (LP) in Eq. (17). Solving (LP) directly is not computationally efficient because there are super-exponentially many variables λ_π. However, we can use the ellipsoid method (Grötschel et al., 2012) to solve (DP) in Eq. (17).

To apply the ellipsoid method to solve (DP), we need to construct a separation oracle, which given any $\gamma \in \mathbb{R}$ and $\mu \in \mathbb{R}_{\geq 0}^n$ such that $\sum_{i \in [n]} x_i \mu_i = 1$, checks whether there is a permutation $\pi \in \mathcal{S}_n$ such that $\sum_{i \in [n]} q_{i,\pi} \mu_i > \gamma$, and if so, outputs the constraint $\sum_{i \in [n]} q_{i,\pi} \mu_i \leq \gamma$. We notice that by Lemma 5.2, $\mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_\pi(A)} \mu_i] \leq \mathbb{E}_{A \sim \mathcal{D}_A}[\sum_{i \in \phi_{\pi_\mu}(A)} \mu_i]$ for all permutations $\pi \in \mathcal{S}_n$, which is equivalent to $\sum_{i \in [n]} q_{i,\pi} \mu_i \leq \sum_{i \in [n]} q_{i,\pi_\mu} \mu_i$ for all $\pi \in \mathcal{S}_n$. Therefore, we can implement the separation oracle efficiently as follows: Given $\gamma \in \mathbb{R}$ and $\mu \in \mathbb{R}_{\geq 0}^n$ such that $\sum_{i \in [n]} x_i \mu_i = 1$, the oracle checks whether $\sum_{i \in [n]} q_{i,\pi_\mu} \mu_i > \gamma$, and if so, outputs the constraint $\sum_{i \in [n]} q_{i,\pi_\mu} \mu_i \leq \gamma$.

Given this separation oracle, we can solve (DP) in Eq. (17) efficiently using the ellipsoid method, which identifies a polynomial number of dual constraints that certify the optimal value of (DP). Then, we can solve (LP) efficiently by restricting it to the variables that correspond to the dual constraints identified by the ellipsoid method. The only issue is that the coefficients x_i and $q_{i,\pi}$ in the LP constraints are not necessarily known. However, this can be addressed by estimating the coefficients *on demand* using Monte-Carlo sampling (i.e., we estimate coefficients x_i for all $i \in [n]$ at the beginning, and we estimate coefficients q_{i,π_μ} for all $i \in [n]$ only when the ellipsoid method queries the separation oracle with input $\mu \in \mathbb{R}_{\geq 0}^n$ and certain $\gamma \in \mathbb{R}$) and solving the approximate versions of the LPs. We defer the details to Section E and state the guarantee of the resulting OCRS in Theorem 5.4.

Theorem 5.4. *For any $\alpha, \varepsilon \in (0, 1]$, there exists an $(\alpha, (1 - \varepsilon)\alpha)$-universal OCRS with preselected order, which given input matroid $\mathcal{M} \subseteq 2^{[n]}$ and α-uncontentious prior distribution $\mathcal{D}_A$ for $\mathcal{M}$, runs in $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}})\right)$ time, where $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$, and $t_{\mathcal{D}_A}, t_{\mathcal{M}}$ are the time it takes to generate a sample from $\mathcal{D}_A$ and to check whether a set of elements belongs to $\mathcal{M}$ respectively.*

6 From secretary algorithms to universal OCRSs (efficiently)

In this section, we show how to use linear programming to efficiently compute a universal OCRS for any arrival model, given a constant-competitive matroid secretary algorithm for that model. Briefly, in the *matroid secretary problem*, we are given a matroid $\mathcal{M} \subseteq 2^{[n]}$ and a weight vector $w \in \mathbb{R}_{\geq 0}^n$. At the beginning, a matroid secretary algorithm knows[6] only $\mathcal{M}$ but not w. Then, elements in $[n]$ arrive in a certain order according to the arrival model. Upon the arrival of each element $i \in [n]$, its weight w_i is revealed, and the algorithm must decide immediately and irrevocably whether to select the element. The goal of the algorithm is to select a set of elements $X \in \mathcal{M}$ with maximum total weight. We say that the algorithm is c-competitive if for any input matroid $\mathcal{M} \subseteq 2^{[n]}$ and

---

[6]We note that many matroid secretary algorithms in the literature do not require full knowledge of $\mathcal{M}$ from the outset. Our result in this section also applies to these algorithms.

weight vector $w \in \mathbb{R}_{\geq 0}^n$, it guarantees that $\mathbb{E}[\sum_{i \in X} w_i] \geq c \cdot r_{\mathcal{M},w}([n])$, where the expectation is taken over the randomness of the algorithm (and possibly the arrival model).

Given a matroid secretary algorithm ALG in any arrival model (we assume w.l.o.g. that ALG only selects elements with strictly positive weights), for each weight vector $w \in \mathbb{R}_{\geq 0}^n$, we construct an OCRS $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ in the same arrival model as ALG: Given input matroid $\mathcal{M} \subseteq 2^{[n]}$ and a set of active elements $A \subseteq [n]$, OCRS $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ provides $\mathcal{M}$ as the input matroid to algorithm ALG. Suppose that elements in $[n]$ arrive in an order $\pi : [n] \to [n]$ according to the arrival model of ALG (it is possible that π is chosen randomly and adaptively by ALG). For each $i \in [n]$, when element $\pi(i)$ arrives, OCRS $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ checks whether $\pi(i) \in A$. If $\pi(i) \in A$, it presents element $\pi(i)$ with weight $w_{\pi(i)}$ to algorithm ALG; otherwise it presents $\pi(i)$ with weight 0 to ALG. OCRS $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ selects element $\pi(i)$ if and only if algorithm ALG selects $\pi(i)$.

This OCRS was originally constructed in the proof of Dughmi (2020, Theorem 4.1). We state its key property in the following lemma.

Lemma 6.1 (Dughmi (2020, Lemma 4.3)). *For any $\alpha, c \in [0,1]$, if the matroid secretary algorithm* ALG *is c-competitive, then given any input matroid $\mathcal{M} \subseteq 2^{[n]}$ and α-uncontentious prior distribution $\mathcal{D}_A$ for $\mathcal{M}$, for any weight vector $w \in \mathbb{R}_{\geq 0}^n$, OCRS $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ guarantees that*

$$\mathbb{E}_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi^{(\mathsf{ALG},w)}} \left[\sum_{i \in \phi(A)} w_i \right] \geq c \cdot \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A} \left[\sum_{i \in A} w_i \right].$$

6.1 Formulating the LP

Now we formulate an LP that given a c-competitive matroid secretary algorithm ALG, computes a nearly $(c \cdot \alpha)$-balanced OCRS for any matroid $\mathcal{M}$ and any α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$. The resulting OCRS will be a random mixture of OCRSs $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ for various weight vectors $w \in \mathbb{R}_{\geq 0}^n$. Specifically, we let $x_i := \Pr_{A \sim \mathcal{D}_A}[i \in A]$ and $q_{i,w} := \Pr_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi^{(\mathsf{ALG},w)}}[i \in \phi(A)]$ for all $i \in [n]$ and $w \in \mathbb{R}_{\geq 0}^n$, and we define $W_\varepsilon := \left\{ \frac{\varepsilon \cdot i}{n} \;\middle|\; i \in \left\{ 0, \ldots, \left\lceil \frac{n}{\varepsilon \cdot p_{\min}} \right\rceil \right\} \right\}$, where $\varepsilon > 0$ is a parameter which we can choose arbitrarily, and $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$ (we assume that $p_{\min} > 0$ as in the preliminary). We consider the following LP (LP1) and its dual (DP1).

$$
\begin{aligned}
\text{(LP1)} \quad & \max_{\beta, \lambda_w} \beta \\
& \text{s.t.} \sum_{w \in W_\varepsilon^n} q_{i,w} \lambda_w \geq \beta x_i \quad \forall i \in [n] \\
& \qquad \sum_{w \in W_\varepsilon^n} \lambda_w = 1 \\
& \qquad \lambda_w \geq 0 \qquad\qquad \forall w \in W_\varepsilon^n. \\
\text{(DP1)} \quad & \min_{\gamma, \mu_i} \gamma \\
& \text{s.t.} \sum_{i \in [n]} q_{i,w} \mu_i \leq \gamma \qquad \forall w \in W_\varepsilon^n \\
& \qquad \sum_{i \in [n]} x_i \mu_i = 1 \\
& \qquad \mu_i \geq 0 \qquad\qquad \forall i \in [n]. \quad\quad (18)
\end{aligned}
$$

We note that every feasible solution (β, λ_w) to (LP1) corresponds to a β-balanced OCRS for $(\mathcal{M}, \mathcal{D}_A)$ in the same arrival model as algorithm ALG. Indeed, because of the last two constraints in (LP1), variables λ_w for all $w \in W_\varepsilon^n$ together specify a distribution $\mathcal{D}_w$ over W_ε^n. We observe that $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ with $w \sim \mathcal{D}_w$ is an OCRS for $(\mathcal{M}, \mathcal{D}_A)$ in the same arrival model as algorithm ALG, and the first constraint in (LP1) ensures that this OCRS is β-balanced.

Next, we show that if the matroid secretary algorithm ALG is c-competitive, and the prior distribution $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, then the optimal value of (DP1) is at least $(1-\varepsilon)c \cdot \alpha$. By LP duality, the optimal value of (LP1) is also at least $(1-\varepsilon)c \cdot \alpha$, and hence, the optimal solution to (LP1) corresponds to an $((1-\varepsilon)c \cdot \alpha)$-balanced OCRS for $(\mathcal{M}, \mathcal{D}_A)$ in the same arrival model as algorithm ALG.

Lemma 6.2. *For any $\alpha, c, \varepsilon \in (0,1]$, if the matroid secretary algorithm* ALG *is c-competitive, then for any matroid $\mathcal{M} \subseteq 2^{[n]}$ and α-uncontentious prior distribution $\mathcal{D}_A$ for $\mathcal{M}$, any vector $\mu \in \mathbb{R}^n$ that satisfies the last two constraints in (DP1) must also satisfy that $\sum_{i \in [n]} q_{i,\mu'} \mu_i \geq (1-\varepsilon)c \cdot \alpha$, where $\mu' \in W_\varepsilon^n$ is defined as follows:*

$$\mu_i' = \max\{y \in W_\varepsilon \mid y \leq \mu_i\} \text{ for all } i \in [n]. \tag{19}$$

This implies that the optimal value of (DP1) is at least $(1-\varepsilon)c \cdot \alpha$.

Proof. Consider any vector $\mu \in \mathbb{R}^n$ that satisfies the last two constraints in (DP1). Let $\mu' \in W_\varepsilon^n$ be the corresponding vector defined by Eq. (19). First, we notice that by Eq. (19), we have that $\mu_i' \leq \mu_i$ for all $i \in [n]$, and it follows that

$$\sum_{i \in [n]} q_{i,\mu'} \mu_i \geq \sum_{i \in [n]} q_{i,\mu'} \mu_i'. \tag{20}$$

Moreover, we show that $\mu_i' \geq \mu_i - \frac{\varepsilon}{n}$ for all $i \in [n]$. Specifically, the last two constraints in (DP1) imply that $\mu_i \leq \frac{1}{x_i} \leq \frac{1}{p_{\min}}$. Notice that the largest number in W_ε is at least $\frac{1}{p_{\min}} \geq \mu_i$, and hence, there exists $j \in \left\{0, \ldots, \left\lceil \frac{n}{\varepsilon \cdot p_{\min}} \right\rceil \right\}$ such that $\frac{\varepsilon \cdot j}{n} \leq \mu_i < \frac{\varepsilon(j+1)}{n}$. It follows that $\mu_i' = \frac{\varepsilon \cdot j}{n} \geq \mu_i - \frac{\varepsilon}{n}$ by Eq. (19). Since $\mu_i' \geq \mu_i - \frac{\varepsilon}{n}$ for all $i \in [n]$, we have that

$$\sum_{i \in [n]} x_i \mu_i' \geq \sum_{i \in [n]} x_i \mu_i - \sum_{i \in [n]} \frac{\varepsilon x_i}{n} \geq \sum_{i \in [n]} x_i \mu_i - \sum_{i \in [n]} \frac{\varepsilon}{n} = \sum_{i \in [n]} x_i \mu_i - \varepsilon = 1 - \varepsilon, \tag{21}$$

where the last inequality is because $x_i \leq 1$ for all $i \in [n]$, and the last equality follows from the second constraint of (DP1). Furthermore, by Lemma 6.1, we have that $\mathbb{E}_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi^{(\mathsf{ALG}, \mu')}}\left[\sum_{i \in \phi(A)} \mu_i'\right] \geq c \cdot \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A}\left[\sum_{i \in A} \mu_i'\right]$, which is equivalent to

$$\sum_{i \in [n]} q_{i,\mu'} \mu_i' \geq c \cdot \alpha \cdot \sum_{i \in [n]} x_i \mu_i'. \tag{22}$$

Combining Ineq. (20)-(22), we get $\sum_{i \in [n]} q_{i,\mu'} \mu_i \geq \sum_{i \in [n]} q_{i,\mu'} \mu_i' \geq c \cdot \alpha \cdot \sum_{i \in [n]} x_i \mu_i' \geq (1-\varepsilon)c \cdot \alpha$. This implies that any feasible solution (γ, μ_i) to (DP1) must satisfy $\gamma \geq (1-\varepsilon)c \cdot \alpha$ because of the constraint $\sum_{i \in [n]} q_{i,\mu'} \mu_i \leq \gamma$. It follows that the optimal value of (DP1) is at least $(1-\varepsilon)c \cdot \alpha$. $\square$

6.2 Reducing the number of dual constraints

We will not directly solve (DP1) in Eq. (18) using the ellipsoid method (because here we do not have a straightforward implementation of the separation oracle). Instead, we will use the ellipsoid

method to reduce the number of constraints in (DP1) such that its optimal value remains at least $(1 - 2\varepsilon)c \cdot \alpha$ (this technique was also used by Lee and Singla (2018) to compute optimal OCRSs for product distributions). Specifically, we consider the following polytope Q_ε:

$$Q_\varepsilon := \{\mu \in \mathbb{R}_{\geq 0}^n \mid \sum_{i \in [n]} x_i \mu_i = 1, \sum_{i \in [n]} q_{i,w} \mu_i \leq (1 - 2\varepsilon)c \cdot \alpha \text{ for all } w \in W_\varepsilon^n\}.$$

If $\alpha, c, \varepsilon \in (0, 1]$, then by Lemma 6.2, any vector $\mu \in \mathbb{R}_{\geq 0}^n$ such that $\sum_{i \in [n]} x_i \mu_i = 1$ must satisfy that $\sum_{i \in [n]} q_{i,\mu'} \mu_i \geq (1 - \varepsilon)c \cdot \alpha$, where $\mu' \in W_\varepsilon^n$ is defined in Eq. (19). Therefore, polytope Q_ε is empty, and moreover, we can construct an efficient separation oracle for Q_ε as follows: Given any $\mu \in \mathbb{R}_{\geq 0}^n$ such that $\sum_{i \in [n]} x_i \mu_i = 1$, the oracle outputs the violated constraint $\sum_{i \in [n]} q_{i,\mu'} \mu_i \leq (1 - 2\varepsilon)c \cdot \alpha$ for $\mu' \in W_\varepsilon^n$ given by Eq. (19). Using this separation oracle, we can apply the ellipsoid method to identify a polynomial-size subset $W' \subseteq W_\varepsilon^n$ such that the following polytope Q'_ε is empty:

$$Q'_\varepsilon := \{\mu \in \mathbb{R}_{\geq 0}^n \mid \sum_{i \in [n]} x_i \mu_i = 1, \sum_{i \in [n]} q_{i,w} \mu_i \leq (1 - 2\varepsilon)c \cdot \alpha \text{ for all } w \in W'\}.$$

Now we consider the following LP (LP2) and its dual (DP2), which are the reduced versions of (LP1) and (DP1) respectively:

$$
\begin{aligned}
\text{(LP2)} \qquad & \max_{\beta, \lambda_w} \beta \\
& \text{s.t.} \quad \sum_{w \in W'} q_{i,w} \lambda_w \geq \beta x_i \quad \forall i \in [n] \\
& \qquad \sum_{w \in W'} \lambda_w = 1 \\
& \qquad \lambda_w \geq 0 \qquad\qquad \forall w \in W'. \\
\text{(DP2)} \qquad & \min_{\gamma, \mu_i} \gamma \\
& \text{s.t.} \quad \sum_{i \in [n]} q_{i,w} \mu_i \leq \gamma \qquad \forall w \in W' \\
& \qquad \sum_{i \in [n]} x_i \mu_i = 1 \\
& \qquad \mu_i \geq 0 \qquad\qquad \forall i \in [n]. \qquad\qquad (23)
\end{aligned}
$$

Because polytope Q'_ε is empty, the optimal value of (DP2) is greater than $(1 - 2\varepsilon)c \cdot \alpha$, and by LP duality, the optimal value of (LP2) is also greater than $(1 - 2\varepsilon)c \cdot \alpha$.

Furthermore, note that (LP2) has polynomially many variables and constraints, and hence, its optimal solution, which we denote by (β^*, λ_w^*), can be computed in polynomial time. By the last two constraints in (LP2), variables λ_w^* for all $w \in W'$ together specify a distribution $\mathcal{D}_w^*$ over W'. We observe that $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ with $w \sim \mathcal{D}_w^*$ is an OCRS for $(\mathcal{M}, \mathcal{D}_A)$ in the same arrival model as the matroid secretary algorithm ALG. This OCRS is β^*-balanced because of the first constraint in (LP2). Thus, it is $((1 - 2\varepsilon)c \cdot \alpha)$-balanced since $\beta^* > (1 - 2\varepsilon)c \cdot \alpha$.

Finally, similar to Subsection 5.2, here we also need to estimate the coefficients x_i and $q_{i,w}$ using Monte-Carlo sampling because they are not necessarily known. We defer the details to Section E and summarize the result in Theorem 6.3.

Theorem 6.3. *For any $\alpha, c, \varepsilon \in (0,1]$, for any arrival model, if there is a c-competitive matroid secretary algorithm* ALG, *then there exists an $(\alpha, (1-\varepsilon)c \cdot \alpha)$-universal OCRS, which given input matroid $\mathcal{M} \subseteq 2^{[n]}$ and α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, runs in $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathsf{ALG}} + t_{\mathcal{D}_A})\right)$ time, where $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$, and t_{ALG} is the worst-case runtime of algorithm* ALG *on matroid secretary problem instances specified by matroid $\mathcal{M}$ and weight vectors $w \in W_\varepsilon^n$, and $t_{\mathcal{D}_A}$ is the time it takes to generate a sample from $\mathcal{D}_A$.*

References

M. Adamczyk and M. Włodarczyk. Random order contention resolution schemes. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 790–801. IEEE, 2018.

S. Agrawal, J. Sethuraman, and X. Zhang. On optimal ordering in the optimal stopping problem. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 187–188, 2020.

M. Babaioff, N. Immorlica, and R. Kleinberg. Matroids, secretary problems, and online mechanisms. In *Symposium on Discrete Algorithms (SODA'07)*, pages 434–443, 2007.

H. Beyhaghi, N. Golrezaei, R. P. Leme, M. Pal, and B. Sivan. Improved approximations for free-order prophets and second-price auctions. *arXiv preprint arXiv:1807.03435*, 2018.

A. Bubna and A. Chiplunkar. Prophet inequality: Order selection beats random order. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 302–336, 2023.

S. Chawla, J. D. Hartline, D. L. Malec, and B. Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 311–320, 2010.

C. Chekuri, J. Vondrák, and R. Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. *SIAM Journal on Computing*, 43(6):1831–1879, 2014.

H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, pages 493–507, 1952.

S. Dughmi. The outer limits of contention resolution on matroids and connections to the secretary problem. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2020.

S. Dughmi. Matroid secretary is equivalent to contention resolution. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022.

S. Dughmi, Y. H. Kalayci, and N. Patel. Limitations of stochastic selection problems with pairwise independent priors. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, page 479–490, 2024.

M. Feldman, O. Svensson, and R. Zenklusen. Online contention resolution schemes with applications to bayesian selection problems. *SIAM Journal on Computing*, 50(2):255–300, 2021.

H. Fu, P. Lu, Z. G. Tang, H. Wu, J. Wu, and Q. Zhang. Sample-based matroid prophet inequalities. In *Proceedings of the 25th ACM Conference on Economics and Computation*, 2024.

M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012.

A. Gupta and V. Nagarajan. A stochastic probing problem with applications. In *Integer Programming and Combinatorial Optimization: 16th International Conference, IPCO 2013, Valparaíso, Chile, March 18-20, 2013. Proceedings 16*, pages 205–216. Springer, 2013.

A. Gupta, J. Hu, G. Kehne, and R. Levin. Pairwise-independent contention resolution. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 196–209. Springer, 2024.

M. T. Hajiaghayi, D. R. Kowalski, P. Krysta, and J. Olkowski. Optimal algorithms for free order multiple-choice secretary. *arXiv preprint arXiv:2207.10703*, 2022.

T. Hill. Prophet inequalities and order selection in optimal stopping problems. *Proceedings of the American Mathematical Society*, 88(1):131–137, 1983.

W. Hoeffding. Probability inequalities for sums of bounded random variables. In *The Collected Works of Wassily Hoeffding*, pages 409–426. Springer, 1994.

P. Jaillet, J. A. Soto, and R. Zenklusen. Advances on matroid secretary problems: Free order model and laminar case. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 254–265. Springer, 2013.

E. Lee and S. Singla. Optimal online contention resolution schemes via ex-ante prophet inequalities. In *26th Annual European Symposium on Algorithms (ESA 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

A. Liu, R. P. Leme, M. Pál, J. Schneider, and B. Sivan. Variable decomposition for prophet inequalities and optimal ordering. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 692–692, 2021.

B. Peng and Z. G. Tang. Order selection prophet inequality: From threshold optimization to arrival time design. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 171–178. IEEE, 2022.

D. J. Welsh. *Matroid theory*. Courier Corporation, 2010.

A   Useful lemmata

Lemma A.1 (see e.g., Welsh (2010))**.** *Any matroid $\mathcal{M} \subseteq 2^{[n]}$ satisfies the following properties:*

    i. For any $X \subseteq [n]$, $r_{\mathcal{M}}(X) \leq |X|$.

    ii. For any $Y \subseteq X \subseteq [n]$ and any weight vector $w \in \mathbb{R}_{\geq 0}^n$, $r_{\mathcal{M},w}(Y) \leq r_{\mathcal{M},w}(X)$.

    iii. For any $X, Y \subseteq [n]$, $r_{\mathcal{M}}(X \mid Y) \leq \sum_{i \in X} r_{\mathcal{M}}(\{i\} \mid Y)$.

    iv. For any $Y \subseteq X \subseteq [n]$, $\operatorname{span}_{\mathcal{M}}(Y) \subseteq \operatorname{span}_{\mathcal{M}}(X)$.

v. *For any $X \subseteq [n]$ and any basis Y of X, an element $i \in [n]$ is not spanned by X if and only if $i \notin Y$ and $Y \cup \{i\} \in \mathcal{M}$.*

vi. *For any $X \subseteq [n]$, the restriction $\mathcal{M}_X$ is also a matroid, and moreover, for any $Y \subseteq X$ and $i \in X$, $i \in \mathrm{span}_{\mathcal{M}_X}(Y)$ holds if and only if $i \in \mathrm{span}_{\mathcal{M}}(Y)$.*

Lemma A.2 (concentration inequalities (Chernoff, 1952; Hoeffding, 1994)). *Let $X = \sum_{i=1}^{n} X_i$ where X_i's are independent random variables taking values in $\{0,1\}$. Then, it holds that*

i. $\Pr[|X - \mathbb{E}[X]| \geq \delta \mathbb{E}[X]] \leq 2 \exp\left(-\frac{\delta^2 \mathbb{E}[X]}{3}\right)$ *for all $\delta \in (0,1)$, and*

ii. $\Pr[|X - \mathbb{E}[X]| \geq t] \leq 2 \exp\left(-\frac{2t^2}{n}\right)$.

B Supplementary Proofs

Lemma B.1 (Restatement of Lemma 2.6). *Given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, for any $S \subseteq [n]$, the marginal distribution $\mathcal{D}_A^S$ is α-uncontentious for $\mathcal{M}_S$.*

Proof. Since $\mathcal{D}_A$ is α-uncontentious for $\mathcal{M}$, there exists a CRS $\mathcal{D}_\phi$ that is α-balanced for $(\mathcal{M}, \mathcal{D}_A)$. For any $S \subseteq [n]$, consider a new CRS $\mathcal{D}_\phi^S$ for $(\mathcal{M}_S, \mathcal{D}_A^S)$, which given an input set of active elements $Z \subseteq S$, outputs a random subset of Z as follows:

(1) Sample a map ϕ from $\mathcal{D}_\phi$, and sample a set A from $\mathcal{D}_A$ conditioned on $A \cap S = Z$.

(2) Then, output the subset $\phi(A) \cap Z$ (which is in $\mathcal{M}_S$ because $\phi(A) \in \mathcal{M}$ and $Z \subseteq S$).

Now we show that $\mathcal{D}_\phi^S$ is an α-balanced CRS for $(\mathcal{M}_S, \mathcal{D}_A^S)$, which will finish the proof. To this end, we derive that for all $i \in S$,

$$\Pr_{A' \sim \mathcal{D}_A^S, \phi' \sim \mathcal{D}_\phi^S}[i \in \phi'(A') \mid i \in A']$$

$$= \sum_{\substack{Z \subseteq S \\ \text{s.t. } i \in Z}} \Pr_{A' \sim \mathcal{D}_A^S, \phi' \sim \mathcal{D}_\phi^S}[A' = Z, i \in \phi'(Z) \mid i \in A']$$

$$= \sum_{\substack{Z \subseteq S \\ \text{s.t. } i \in Z}} \Pr_{A' \sim \mathcal{D}_A^S}[A' = Z \mid i \in A'] \times \Pr_{\phi' \sim \mathcal{D}_\phi^S}[i \in \phi'(Z)] \qquad (A', \phi' \text{ are independent})$$

$$= \sum_{\substack{Z \subseteq S \\ \text{s.t. } i \in Z}} \Pr_{A' \sim \mathcal{D}_A^S}[A' = Z \mid i \in A'] \times \Pr_{\substack{A \sim \mathcal{D}_A, \\ \phi \sim \mathcal{D}_\phi}}[i \in \phi(A) \cap Z \mid A \cap S = Z] \qquad (\text{By definition of } \mathcal{D}_\phi^S)$$

$$= \sum_{\substack{Z \subseteq S \\ \text{s.t. } i \in Z}} \Pr_{A \sim \mathcal{D}_A}[A \cap S = Z \mid i \in A \cap S] \times \Pr_{\substack{A \sim \mathcal{D}_A, \\ \phi \sim \mathcal{D}_\phi}}[i \in \phi(A) \cap Z \mid A \cap S = Z] \quad (\text{By definition of } \mathcal{D}_A^S)$$

$$= \sum_{\substack{Z \subseteq S \\ \text{s.t. } i \in Z}} \Pr_{A \sim \mathcal{D}_A}[A \cap S = Z \mid i \in A \cap S] \times \Pr_{\substack{A \sim \mathcal{D}_A, \\ \phi \sim \mathcal{D}_\phi}}[i \in \phi(A) \mid A \cap S = Z, i \in A \cap S]$$

$$\qquad (\text{Since } i \in Z)$$

$$= \sum_{\substack{Z \subseteq S \\ \text{s.t. } i \in Z}} \Pr_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}[i \in \phi(A), A \cap S = Z \mid i \in A \cap S]$$

$$= \Pr_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}[i \in \phi(A) \mid i \in A \cap S]$$

$$= \Pr_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}[i \in \phi(A) \mid i \in A] \qquad\qquad \text{(Since } i \in S\text{)},$$

which is at least α because $\mathcal{D}_\phi$ is α-balanced for $(\mathcal{M}, \mathcal{D}_A)$. $\qquad\qquad \square$

Proposition B.1. *Consider the 1-uniform matroid $\mathcal{M}$ and the prior distributions $\mathcal{D}_A^{(j)}$ for all $j \in [n]$, which are defined in Example 2.7. Suppose that given instance $(\mathcal{M}, \mathcal{D}_A^{(j)})$ for any $j \in [n]$, an algorithm ALG only has sample access to $\mathcal{D}_A^{(j)}$, and it draws at most $o(\frac{1}{\delta})$ samples from $\mathcal{D}_A^{(j)}$. Then, there exists $j \in [n]$ such that ALG is at most $o(1)$-balanced for $(\mathcal{M}, \mathcal{D}_A^{(j)})$.*

Proof. By the assumption in the proposition, given instance $(\mathcal{M}, \mathcal{D}_A^{(j)})$ for any $j \in [n]$, ALG draws at most $m = o(\frac{1}{\delta})$ samples from $\mathcal{D}_A^{(j)}$. We consider the following random process:

(1) Sample $j^* \in [n]$ uniformly at random.

(2) Sample m sets $A_1, \ldots, A_m \subseteq [n]$ from $\mathcal{D}_A^{(j^*)}$.

(3) Given access to matroid $\mathcal{M}$ and samples $A_1, \ldots, A_m$ from $\mathcal{D}_A^{(j^*)}$, run ALG with the input set of active elements $A = [n]$. Let $j' \in [n]$ denote the element that is selected by ALG (we assume w.l.o.g. that ALG always selects an element when the input set of active elements is $[n]$).

Suppose that ALG is β-balanced for $(\mathcal{M}, \mathcal{D}_A^{(j)})$ for all $j \in [n]$. Then, given instance $(\mathcal{M}, \mathcal{D}_A^{(j)})$ for any $j \in [n]$, ALG must select element j with probability at least β if the input set of active elements is $[n]$, because this is the only case where element j is active by our construction of $\mathcal{D}_A^{(j)}$. This implies that $\Pr[j' = j^*] \geq \beta$ in the above random process. Therefore, to prove the proposition, it suffices to show that $\Pr[j' = j^*] = o(1)$.

To this end, we let $J(A_1, \ldots, A_m) := \{j \in [n] \mid \forall \ell \in [m], A_\ell \neq \{j\}\}$. By our construction of the prior distributions $\mathcal{D}_A^{(j)}$ for all $j \in [n]$, we have that for any $j \notin J(A_1, \ldots, A_m)$, the likelihood of observing $A_1, \ldots, A_m$ in the above random process is 0 if $j^* = j$, and for any $j_1, j_2 \in J(A_1, \ldots, A_m)$, the likelihood of observing $A_1, \ldots, A_m$ in the above process is the same regardless of whether $j^* = j_1$ or $j^* = j_2$, namely, $\Pr[A_1, \ldots, A_m \mid j^* = j_1] = \Pr[A_1, \ldots, A_m \mid j^* = j_2]$. Moreover, since j^* is chosen uniformly at random from $[n]$, we have that $\Pr[j^* = j_1] = \Pr[j^* = j_2]$. It follows that

$$\begin{aligned} \Pr[j^* = j_1 \mid A_1, \ldots, A_m] &= \frac{\Pr[A_1, \ldots, A_m \mid j^* = j_1] \times \Pr[j^* = j_1]}{\Pr[A_1, \ldots, A_m]} \\ &= \frac{\Pr[A_1, \ldots, A_m \mid j^* = j_2] \times \Pr[j^* = j_2]}{\Pr[A_1, \ldots, A_m]} \\ &= \Pr[j^* = j_2 \mid A_1, \ldots, A_m]. \end{aligned}$$

Thus, conditioned on $A_1, \ldots, A_m$, j^* is a uniformly random number in $J(A_1, \ldots, A_m)$. Moreover, we observe that conditioned on $A_1, \ldots, A_m$, j' is independent of j^*. Therefore, we have that

$$\Pr[j' = j^* \mid A_1, \ldots, A_m] \leq \frac{1}{|J(A_1, \ldots, A_m)|}. \tag{24}$$

Next, we show that $|J(A_1, \ldots, A_m)|$ is large with high probability. We first notice that for all $j \in [n]$, by our construction of $\mathcal{D}_A^{(j)}$, we have that $\Pr_{A \sim \mathcal{D}_A^{(j)}}[A = \{i\}] = \delta$ for any $i \in [n]$ such that $i \neq j$. Hence, in the above random process, we have that for all $j \in [n]$ and $\ell \in [m]$,

$$\Pr[A_\ell = \{i\} \mid j^* = j] = \Pr_{A_\ell \sim \mathcal{D}_A^{(j)}}[A_\ell = \{i\}] = \delta \text{ for any } i \in [n] \text{ such that } i \neq j.$$

It follows that for any $i, j \in [n]$ such that $i \neq j$,

$$\Pr[i \in J(A_1, \ldots, A_m) \mid j^* = j] = \prod_{\ell \in [m]} \Pr[A_\ell \neq \{i\} \mid j^* = j] = (1 - \delta)^m = (1 - \delta)^{o(\frac{1}{\delta})} = 1 - o(1),$$

which implies that for any $j \in [n]$,

$$\mathbb{E}[|J(A_1, \ldots, A_m)| \mid j^* = j] = \sum_{i \in [n] \text{ s.t. } i \neq j} \Pr[i \in J(A_1, \ldots, A_m) \mid j^* = j] = (1 - o(1)) \cdot (n - 1).$$

Therefore, we have that $\mathbb{E}[|J(A_1, \ldots, A_m)|] = (1 - o(1))n$, and by Markov's inequality, we have that $\Pr[n - |J(A_1, \ldots, A_m)| \geq \frac{n}{2}] \leq \frac{n - \mathbb{E}[|J(A_1, \ldots, A_m)|]}{n/2} = \frac{o(n)}{n/2} = o(1)$, which is equivalent to $\Pr[|J(A_1, \ldots, A_m)| \leq \frac{n}{2}] = o(1)$. We finish the proof by noticing that

$$\Pr[j' = j^*] \leq \Pr\left[j' = j^* \,\middle|\, |J(A_1, \ldots, A_m)| \geq \frac{n}{2}\right] + \Pr\left[|J(A_1, \ldots, A_m)| \leq \frac{n}{2}\right]$$

$$\leq \Pr\left[j' = j^* \,\middle|\, |J(A_1, \ldots, A_m)| \geq \frac{n}{2}\right] + o(1)$$

$$\leq \frac{2}{n} + o(1) = o(1) \qquad\qquad\qquad\qquad\qquad \text{(By Eq. (24)).}$$

$\square$

Fact B.2 (Restatement of Fact 4.5). *For any integer $i \geq 1$, for any real numbers $\alpha, x_0, \ldots, x_{i-1} \in [0, 1]$, if $\sum_{k=0}^{i-1} \frac{1}{i} \cdot x_k \geq \alpha$, then $\sum_{k=0}^{i-1} \frac{k+1}{i(i+1)} \cdot x_k \geq \frac{\alpha^2}{2}$.*

Proof. To prove this fact, it suffices to show that for any $\alpha \in [0, 1]$, $\frac{\alpha^2}{2}$ is a lower bound of the optimal value of the following LP,

$$\min_{x_0, \ldots, x_{i-1}} \sum_{k=0}^{i-1} \frac{k+1}{i(i+1)} \cdot x_k$$

$$\text{s.t.} \sum_{k=0}^{i-1} x_k \geq \alpha \cdot i$$

$$\forall j \in \{0, \ldots, i-1\}, \ 0 \leq x_j \leq 1. \tag{25}$$

Intuitively, the LP in Eq. (25) is the "opposite" of the fractional knapsack problem.

Deriving the optimal solution to the LP

First, we observe that the optimal solution to the LP satisfies $\sum_{k=0}^{i-1} x_k = \alpha \cdot i$, because otherwise we can decrease any strictly positive variable x_k in the optimal solution slightly without violating the LP constraints, which decreases the objective value. Moreover, since the coefficient of variable x_k in the objective strictly increases as k increases, for any $k_1, k_2 \in \{0, 1, \ldots, i-1\}$ such that $k_1 < k_2$, the optimal solution should assign as much value as possible to variable x_{k_1} before it assigns any value to variable x_{k_2}. Therefore, the optimal solution to the LP is the following:

$$x_k^* = \begin{cases} 1 & \text{for } 0 \leq k < m \\ r & \text{for } k = m \\ 0 & \text{for } m < k \leq i - 1, \end{cases} \tag{26}$$

where $m := \lfloor \alpha \cdot i \rfloor$ and $r := \alpha \cdot i - \lfloor \alpha \cdot i \rfloor$.

Lower bounding the objective value of the optimal solution

Next, we prove that the objective value of the optimal solution $(x_0^*, \ldots, x_{i-1}^*)$, which we denote by OPT, is at least $\frac{\alpha^2}{2}$. We divide the proof into two cases.

Case 1: $\alpha \cdot i \geq 1$

If $\alpha \cdot i \geq 1$, we have that $m = \lfloor \alpha \cdot i \rfloor \geq 1$, and we can lower bound OPT as follows,

$$
\begin{aligned}
\text{OPT} &= \sum_{k=0}^{i-1} \frac{k+1}{i(i+1)} \cdot x_k^* = \sum_{k=0}^{m-1} \frac{k+1}{i(i+1)} \cdot 1 + \frac{m+1}{i(i+1)} \cdot r && \text{(By Eq. (26))} \\
&= \frac{m(m+1)}{2i(i+1)} + \frac{r(m+1)}{i(i+1)} = \frac{(m+2r)(m+1)}{2i(i+1)} \\
&= \frac{(m+r+r)(m+r+1-r)}{2i(i+1)} = \frac{(\alpha \cdot i + r)(\alpha \cdot i + 1 - r)}{2i(i+1)} && \text{(Since } m + r = \alpha \cdot i) \\
&= \frac{(\alpha \cdot i)^2 + \alpha \cdot i + r(1-r)}{2i(i+1)} \geq \frac{(\alpha \cdot i)^2 + \alpha \cdot i}{2i(i+1)} && \text{(Since } r = \alpha \cdot i - \lfloor \alpha \cdot i \rfloor \in [0,1]) \\
&= \frac{\alpha \cdot i(\alpha \cdot i + 1)}{2i(i+1)} \geq \frac{\alpha \cdot i(\alpha \cdot i + \alpha)}{2i(i+1)} && \text{(Since } \alpha \leq 1) \\
&= \frac{\alpha^2}{2}.
\end{aligned}
$$

Case 2: $\alpha \cdot i < 1$

If $\alpha \cdot i < 1$, we have that $m = \lfloor \alpha \cdot i \rfloor = 0$, and we can lower bound OPT as follows,

$$
\begin{aligned}
\text{OPT} &= \sum_{k=0}^{i-1} \frac{k+1}{i(i+1)} \cdot x_k^* = \frac{1}{i(i+1)} \cdot r && \text{(By Eq. (26))} \\
&= \frac{\alpha \cdot i}{i(i+1)} && \text{(Since } r = \alpha \cdot i - \lfloor \alpha \cdot i \rfloor = \alpha \cdot i \text{ in this case)} \\
&= \frac{\alpha}{i+1} \geq \frac{\alpha}{i+1} \cdot \alpha \cdot i && \text{(Since } \alpha \cdot i < 1) \\
&= \alpha^2 \cdot \left(1 - \frac{1}{i+1}\right) \geq \frac{\alpha^2}{2} && \text{(Since } i \geq 1).
\end{aligned}
$$

$\square$

C   Implementing Algorithm 1 via Monte-Carlo sampling

Both Algorithm 1 and Algorithm 3 preselect the arrival order by iteratively identifying elements that have relatively low probabilities of being spanned by the subsampled active elements (see Line 4 of Algorithm 1 and Line 4 of Subroutine 2). In this section, we demonstrate how to estimate those probabilities using Monte-Carlo sampling and establish bounds on the time and sample complexity for Algorithm 1 (similar bounds can be established for Algorithm 3 analogously).

Corollary C.1 (Restatement of Corollary 3.5)**.** *For any* $\alpha, \varepsilon \in (0,1]$, *there exists an* $(\alpha, \frac{(1-\varepsilon)\alpha^2}{4})$-*universal OCRS with preselected order, which given input matroid* $\mathcal{M} \subseteq 2^{[n]}$ *and* α-*uncontentious*

distribution $\mathcal{D}_A$, runs in $O\left(\frac{n\log(n/\varepsilon)}{\alpha^2\varepsilon^2 p_{\min}} \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}} \cdot n)\right)$ time, where $p_{\min} := \min_{i\in[n]} \Pr_{A\sim\mathcal{D}_A}[i \in A]$, and $t_{\mathcal{D}_A}, t_{\mathcal{M}}$ are the time it takes to generate a sample from $\mathcal{D}_A$ and to check whether a set of elements is in $\mathcal{M}$ respectively.

Proof. We first elaborate how to use Monte-Carlo sampling to implement Line 4 of Algorithm 1, and then we prove the universality guarantee and analyze the total runtime.

Implementing Line 4 of Algorithm 1 through Monte-Carlo sampling

We implement Line 4 of Algorithm 1 using Monte-Carlo sampling as follows:

(1) Draw $m := \left\lceil \frac{128\ln(4n/\varepsilon)}{\alpha^2\varepsilon^2 p_{\min}} \right\rceil$ independent samples $A_1, \ldots, A_m$ from the α-uncontentious prior distribution $\mathcal{D}_A$, and then for each $\ell \in [m]$, sample a subset $B_\ell \subseteq A_\ell$ by including each element of A_ℓ into B_ℓ independently with probability $\frac{\alpha}{2}$.

(2) For each $j \in S_i$, compute $m_j := |\{\ell \in [m] \mid j \in A_\ell\}|$ and $k_j := |\{\ell \in [m] \mid j \notin \text{span}_{\mathcal{M}}(B_\ell \cap S_i) \text{ and } j \in A_\ell\}|$ as follows:

    (i) Initialize $m_j = 0$ and $k_j = 0$ for all $j \in [n]$.

    (ii) For each $\ell \in [m]$, compute a basis[7] X_ℓ of $B_\ell \cap S_i$.

    (iii) For each $\ell \in [m]$ and $j \in S_i$, increase m_j by 1 if $j \in A_\ell$, and increase k_j by 1 if $j \in A_\ell$ and $j \notin X_\ell$ and $X_\ell \cup \{j\} \in \mathcal{M}$ (note that $j \notin X_\ell$ and $X_\ell \cup \{j\} \in \mathcal{M}$ iff $j \notin \text{span}_{\mathcal{M}}(B_\ell \cap S_i)$ by Lemma A.1-v.).

(3) Let $\pi(i)$ be any $j \in S_i$ such that $k_j \geq \frac{(1-\varepsilon/4)\alpha}{2} \cdot m_j$. If such j does not exist, let Algorithm 1 terminate and output an empty set.

To prove that with the above implementation of Line 4, Algorithm 1 is $(\alpha, \frac{(1-\varepsilon)\alpha^2}{4})$-universal, we first establish the following claim.

Claim C.2. *With probability at least $1 - \frac{\varepsilon}{4n}$, the above implementation finds an element $\pi(i)$ which satisfies that $\Pr_{A'\sim\mathcal{D}_A^{S_i}}[\pi(i) \notin \text{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \geq \frac{(1-\varepsilon/2)\alpha}{2}$.*

Proof of Claim C.2. We divide the proof into two steps and then put them together.

Step 1: for all $j \in S_i$, m_j is sufficiently large w.h.p.

Notice that for all $j \in S_i$, $\mathbb{E}[m_j] = \sum_{\ell=1}^m \Pr_{A_\ell\sim\mathcal{D}_A}[j \in A_\ell] \geq p_{\min} \cdot m$, and it follows that

$$
\begin{aligned}
\Pr\left[m_j \leq \frac{p_{\min} \cdot m}{2}\right] &\leq \Pr\left[m_j \leq \frac{\mathbb{E}[m_j]}{2}\right] && \text{(By } \mathbb{E}[m_j] \geq p_{\min} \cdot m\text{)} \\
&\leq 2\exp\left(-\frac{\mathbb{E}[m_j]}{12}\right) && \text{(By Lemma A.2-i.)} \\
&\leq 2\exp\left(-\frac{p_{\min} \cdot m}{12}\right) && \text{(By } \mathbb{E}[m_j] \geq p_{\min} \cdot m\text{)} \\
&\leq 2\exp\left(-\frac{2\ln(4n/\varepsilon)}{\alpha^2\varepsilon^2}\right) && \text{(By our choice of } m\text{)} \\
&\leq 2\exp\left(-2\ln(4n/\varepsilon)\right) = \frac{\varepsilon^2}{8n^2}. && (27)
\end{aligned}
$$

---

[7]This can be done by removing elements from set $B_\ell \cap S_i$ one by one until the set is in $\mathcal{M}$, which takes at most n queries to the membership oracle of $\mathcal{M}$.

Step 2: for all $j \in S_i$, k_j is well-bounded w.h.p.

For each $j \in S_i$ and $\ell \in [m]$, we define random variable $z_j^\ell := \mathbb{1}(j \notin \mathrm{span}_{\mathcal{M}}(B_\ell \cap S_i), j \in A_\ell)$ and let $q_j := \mathrm{Pr}_{A \sim \mathcal{D}_A}[j \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A \cap S_i)) \mid j \in A]$. First, observe that $k_j = \sum_{\ell=1}^m z_j^\ell$. Moreover, notice that conditioned on the event $j \in A_\ell$, z_j^ℓ is a Bernoulli random variable that has value 1 with probability q_j, and conditioned on the event $j \notin A_\ell$, z_j^ℓ is always 0. Hence, for any $L \subseteq [m]$, conditioned on the event that $\{\ell \in [m] \mid j \in A_\ell\} = L$, k_j is a binomial random variable which follows distribution $\mathrm{Bin}(|L|, q_j)$. Therefore, for any $b \in \{0, 1, \ldots, m\}$, conditioned on the event that $|\{\ell \in [m] \mid j \in A_\ell\}| = b$ (in other words, $m_j = b$), k_j is a binomial random variable following distribution $\mathrm{Bin}(b, q_j)$.

Now we show that conditioned on the event $m_j \geq \frac{p_{\min} \cdot m}{2}$, the probability that $k_j \geq \frac{(1-\varepsilon/4)\alpha}{2} \cdot m_j$ is almost zero for any $j \in S_i$ with $q_j \leq \frac{(1-\varepsilon/2)\alpha}{2}$ and is almost one for any $j \in S_i$ with $q_j \geq \frac{\alpha}{2}$. Specifically, because k_j follows distribution $\mathrm{Bin}(b, q_j)$ conditioned on the event $m_j = b$, we have that $\mathbb{E}[k_j \mid m_j = b] = q_j \cdot b$. For any integer $b \geq \frac{p_{\min} \cdot m}{2}$ and any $j \in S_i$ with $q_j \leq \frac{(1-\varepsilon/2)\alpha}{2}$, we derive that

$$\mathrm{Pr}\left[k_j \geq \frac{(1-\varepsilon/4)\alpha \cdot m_j}{2} \,\middle|\, m_j = b\right]$$

$$\leq \mathrm{Pr}\left[k_j \geq \mathbb{E}[k_j \mid m_j = b] + \frac{\varepsilon\alpha \cdot b}{8} \,\middle|\, m_j = b\right] \quad \text{(Since } \mathbb{E}[k_j \mid m_j = b] = q_j \cdot b \leq \frac{(1-\varepsilon/2)\alpha \cdot b}{2})$$

$$\leq 2\exp\left(-\frac{\varepsilon^2\alpha^2 \cdot b}{32}\right) \quad \text{(By Lemma A.2-ii.)}$$

$$\leq 2\exp\left(-\frac{\varepsilon^2\alpha^2 \cdot p_{\min} \cdot m}{64}\right) \quad \text{(Since } b \geq \frac{p_{\min} \cdot m}{2})$$

$$= 2\exp(-2\ln(4n/\varepsilon)) = \frac{\varepsilon^2}{8n^2} \quad \text{(By our choice of } m). \tag{28}$$

Similarly, for any integer $b \geq \frac{p_{\min} \cdot m}{2}$ and any $j \in S_i$ with $q_j \geq \frac{\alpha}{2}$, we have that

$$\mathrm{Pr}\left[k_j \leq \frac{(1-\varepsilon/4)\alpha \cdot m_j}{2} \,\middle|\, m_j = b\right] \leq \frac{\varepsilon^2}{8n^2} \quad \text{(Analogous to Ineq. (28)).} \tag{29}$$

Putting the two steps together

We let E_1 be the event that

$$\forall j \in S_i, \left(q_j \leq \frac{(1-\varepsilon/2)\alpha}{2} \implies k_j < \frac{(1-\varepsilon/4)\alpha \cdot m_j}{2}\right) \wedge \left(q_j \geq \frac{\alpha}{2} \implies k_j > \frac{(1-\varepsilon/4)\alpha \cdot m_j}{2}\right).$$

Since there exists $j \in S_i$ such that $q_j \geq \frac{\alpha}{2}$ (as shown by Ineq. (3) in the proof of Theorem 3.4), event E_1 implies that Step (3) of the above implementation finds an element $\pi(i) = j$ such that $q_j \geq \frac{(1-\varepsilon/2)\alpha}{2}$. Thus, it remains to prove that E_1 happens w.h.p. To this end, let E_2 denote the event that $\forall j \in S_i, m_j \geq \frac{p_{\min} \cdot m}{2}$. By Ineq. (27) and a union bound, we have that $\mathrm{Pr}[\bar{E}_2] \leq \frac{\varepsilon^2}{8n^2} \cdot |S_i| \leq \frac{\varepsilon^2}{8n}$. By Ineq. (28) and (29) and a union bound, we get that $\mathrm{Pr}[\bar{E}_1 \mid E_2] \leq \frac{\varepsilon^2}{8n^2} \cdot |S_i| \leq \frac{\varepsilon^2}{8n}$. It follows that $\mathrm{Pr}[\bar{E}_1] \leq \mathrm{Pr}[\bar{E}_2] + \mathrm{Pr}[\bar{E}_1 \mid E_2] \times \mathrm{Pr}[E_2] \leq \mathrm{Pr}[\bar{E}_2] + \mathrm{Pr}[\bar{E}_1 \mid E_2] \leq \frac{\varepsilon^2}{8n} + \frac{\varepsilon^2}{8n} \leq \frac{\varepsilon}{4n}$. $\qquad \square$

Establishing $(\alpha, \frac{(1-\varepsilon)\alpha^2}{4})$-universality

Claim C.2 shows that with probability at least $1 - \frac{\varepsilon}{4n}$, the above implementation finds an element $\pi(i)$ such that $\mathrm{Pr}_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \geq \frac{(1-\varepsilon/2)\alpha}{2}$, for the i-th iteration of

the first for loop of Algorithm 1. By a union bound, with probability at least $1 - \frac{\varepsilon}{4}$, the above implementation finds element $\pi(i)$ such that $\Pr_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \geq \frac{(1-\varepsilon/2)\alpha}{2}$ for all iterations $i \in [n]$. If Algorithm 1 successfully finds such elements $\pi(i)$'s for all iterations $i \in [n]$, the probability that element $\pi(i)$ is selected in the second for loop of Algorithm 1, conditioned on $\pi(i)$ being active, is at least $\Pr_{A' \sim \mathcal{D}_A^{S_i}}[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(A')) \mid \pi(i) \in A'] \cdot \frac{\alpha}{2} \geq \frac{(1-\varepsilon/2)\alpha^2}{4}$ (as shown by Ineq. (5) in the proof of Theorem 3.4). Therefore, overall, the probability that each element is selected conditioned on it being active is at least $(1 - \frac{\varepsilon}{4}) \cdot \frac{(1-\varepsilon/2)\alpha^2}{4} \geq \frac{(1-\varepsilon)\alpha^2}{4}$.

Analyzing the overall runtime

First, we note that the second for loop of Algorithm 1 requires $O(n)$ queries to the membership oracle of $\mathcal{M}$, which can be done in $O(t_{\mathcal{M}} \cdot n)$ time.

Besides, in the above implementation of Line 4, we sample m sets from $\mathcal{D}_A$ and subsample those sets, which takes $O(t_{\mathcal{D}_A} \cdot m + nm)$ time. Then, we count m_j and k_j for all $j \in S_i$, which requires $O(nm)$ membership queries and hence takes $O(t_{\mathcal{M}} \cdot nm)$ time in total. Thus, the runtime of the above implementation of Line 4 is $O(m(t_{\mathcal{D}_A} + t_{\mathcal{M}} \cdot n))$. It follows that with this implementation, the first for loop of Algorithm 1 takes $O(nm(t_{\mathcal{D}_A} + t_{\mathcal{M}} \cdot n))$ time.

Therefore, the overall runtime is $O(nm(t_{\mathcal{D}_A} + t_{\mathcal{M}} \cdot n)) = O\left(\frac{n \log(n/\varepsilon)}{\alpha^2 \varepsilon^2 p_{\min}} \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}} \cdot n)\right)$. $\qquad\square$

D Tight instances for Algorithm 1 and Algorithm 3

In Theorem 3.4 and Theorem 4.4, we showed that Algorithm 1 and Algorithm 3 are $(\alpha, \Omega(\alpha^2))$-universal respectively; in the analysis, the quadratic exponent in their universality guarantees arose from balancing two conflicting goals which we hope to achieve through subsampling: (i) removing elements such that they do not span other elements, and (ii) retaining elements such that they remain selectable. In this section, we provide concrete examples to demonstrate that the universality guarantees that we established for Algorithm 1 and Algorithm 3 are essentially tight. We first show this for Algorithm 1 using Example D.1.

Example D.1. *For any $\alpha \in \left(0, \frac{1}{2}\right]$ such that $n := \frac{1}{\alpha}$ is an integer, let m be the largest integer such that $\left(1 - \frac{\alpha}{2}\right) \cdot \left(1 - \frac{\alpha^2}{4}\right)^m \geq \frac{\alpha}{2}$ (and hence $\left(1 - \frac{\alpha^2}{4}\right)^m \leq \frac{\alpha}{2(1-\alpha^2/4)(1-\alpha/2)}$). Let $G_1 = (V_1, E_1)$ be the graph in Figure 1a, where $V_1 = \{w, w'\}$, and E_1 consists of n parallel edges $e_1, \ldots, e_n$ between w and w'. Let $G_2 = (V_2, E_2)$ be the hat example in Figure 1b, where $V_2 = \{v_1, \ldots, v_m, u, u'\}$ and $E_2 = \{(v_i, u), (v_i, u') \mid i \in [m]\} \cup \{(u, u')\}$. Let graph $G = (V_1 \cup V_2, E_1 \cup E_2)$ be the union of G_1 and G_2. Then, let $\mathcal{M} \subseteq 2^{E_1 \cup E_2}$ be the graphic matroid[8] of graph G. Let $\mathcal{D}_A \in \Delta(2^{E_1 \cup E_2})$ be the trivial prior distribution such that all edges in $E_1 \cup E_2$ are always active.*

Proposition D.1. *In Example D.1, the prior distribution $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, but Algorithm 1 is at most $O(\alpha^2)$-balanced for $(\mathcal{M}, \mathcal{D}_A)$. Moreover, as α approaches 0, Algorithm 1 is arbitrarily close to being $\frac{\alpha^2}{4}$-balanced for $(\mathcal{M}, \mathcal{D}_A)$.*

Proof. We start by proving that in Example D.1, $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$.

---

[8]Specifically, the graphic matroid of a graph consists of all subsets of edges that do not contain a cycle.
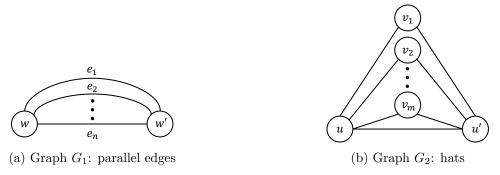
(a) Graph G_1: parallel edges   (b) Graph G_2: hats

Figure 1: Illustration of Example D.1.

$\mathcal{D}_A$ is α-uncontentious for $\mathcal{M}$

We consider the following CRS:

(1) Initialize $X = \{(u, u')\}$.

(2) For each $i \in [m]$, with probability $\frac{1}{2}$, add edge (v_i, u) to X, and with probability $\frac{1}{2}$, add edge (v_i, u') to X.

(3) Select an edge from $\{e_1, \ldots, e_n\}$ uniformly at random and add it to X. Then, output X.

We note that X is always feasible, because it does not contain multiple edges between w and w', and it does not contain a *hat* (i.e., a pair of edges (v_i, u) and (v_i, u') for any $i \in [m]$). Moreover, We observe that for all $i \in [n]$, the probability that edge e_i is included in X is $\frac{1}{n} = \alpha$, and for all $i \in [m]$, the probability that edge (v_i, u) (or (v_i, u')) is included in X is $\frac{1}{2} \geq \alpha$, and the edge (u, u') is always included in X. Therefore, the above CRS is α-balanced for $(\mathcal{M}, \mathcal{D}_A)$, and hence, $\mathcal{D}_A$ is α-uncontentious for $\mathcal{M}$.

Order π satisfies the order-selecting condition of Algorithm 1

Now we show that order π specified by Eq. (30) satisfies the order-selecting condition at Line 4 of Algorithm 1, and then we prove that Algorithm 1 with preselected order π is $O(\alpha^2)$-balanced. Essentially, order π starts with the hats in G_2, followed by the edge at the bottom in G_2, and ends with the parallel edges in G_1.

$$\pi(i) = \begin{cases} (v_i, u) & \text{for } 1 \leq i \leq m \\ (v_{i-m}, u') & \text{for } m+1 \leq i \leq 2m \\ (u, u') & \text{for } i = 2m+1 \\ e_{i-2m-1} & \text{for } 2m+2 \leq i \leq 2m+n+1. \end{cases} \tag{30}$$

For each $i \in [2m]$, element $\pi(i)$ is an edge of a hat in G_2, which cannot be spanned by any set of edges that do not contain $\pi(i)$ and the other edge in the hat (we use the notation $\pi(i)'$ to refer to the other edge in the hat for convenience). Therefore, for all $i \in [2m]$, we have that

$$\Pr[\pi(i) \notin \operatorname{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2))] \geq \Pr[\pi(i) \notin \mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2), \pi(i)' \notin \mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2)]$$

$$= \left(1 - \frac{\alpha}{2}\right)^2 = 1 - \alpha + \frac{\alpha^2}{4} \geq 2\alpha - \alpha + \frac{\alpha^2}{4} \geq \alpha,$$

31

which implies that for all $i \in [2m]$, element $\pi(i)$ satisfies the condition at Line 4 of Algorithm 1. Moreover, element $\pi(2m + 1)$ is edge (u, u') in G_2, which cannot be spanned by any set of edges that do not contain any hat and $\pi(2m + 1)$ itself. It follows that

$$\Pr[\pi(2m + 1) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2))]$$
$$\geq \Pr[\pi(2m + 1) \notin \mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2)] \times \Pr[\forall j \in [m], \{(v_j, u), (v_j, u')\} \notin \mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2)]$$
$$= \left(1 - \frac{\alpha}{2}\right) \cdot \left(1 - \frac{\alpha^2}{4}\right)^m \geq \frac{\alpha}{2} \qquad \text{(By our choice of } m\text{),}$$

which implies that element $\pi(2m+1)$ satisfies the condition at Line 4 of Algorithm 1. Furthermore, for each $i \in \{2m + 2, \ldots, 2m + n + 1\}$, element $\pi(i)$ is an edge in G_1, which cannot be spanned by any set of edges that do not contain $\pi(i)$ and any other edge in G_1, and hence, we have that

$$\Pr[\pi(i) \notin \mathrm{span}_{\mathcal{M}}(\mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2))] \geq \Pr[\forall j \in [n], e_j \notin \mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2)]$$
$$\geq 1 - \sum_{j \in [n]} \Pr[e_j \in \mathcal{T}_{\frac{\alpha}{2}}(E_1 \cup E_2)] \qquad \text{(By a union bound)}$$
$$= 1 - n \cdot \frac{\alpha}{2} = \frac{1}{2} \geq \alpha,$$

which implies that element $\pi(i)$ satisfies the condition at Line 4 of Algorithm 1.

Algorithm 1 is $O(\alpha^2)$-balanced for $(\mathcal{M}, \mathcal{D}_A)$

Suppose that Algorithm 1 uses order π defined by Eq. (30) as the preselected order. We observe that element $\pi(2m + 1)$ (i.e., edge (u, u')) is selected by Algorithm 1 iff (i) $\pi(2m + 1) \in T$, and (ii) for all $i \in [m]$, at least one of the two edges (v_i, u) and (v_i, u') does not belong to T. Thus, the probability that element $\pi(2m + 1)$ is selected by Algorithm 1 is $\frac{\alpha}{2} \cdot \left(1 - \frac{\alpha^2}{4}\right)^m$, which is at most $\frac{\alpha}{2} \cdot \frac{\alpha}{2(1-\alpha^2/4)(1-\alpha/2)}$ by our choice of m. This implies that Algorithm 1 is $O(\alpha^2)$-balanced for $(\mathcal{M}, \mathcal{D}_A)$. Finally, we remark that as α approaches 0, the ratio $\frac{\alpha}{2} \cdot \frac{\alpha}{2(1-\alpha^2/4)(1-\alpha/2)}$ approaches $\frac{\alpha^2}{4}$, and hence, Algorithm 1 is arbitrarily close to being $\frac{\alpha^2}{4}$-balanced for $(\mathcal{M}, \mathcal{D}_A)$. $\qquad \square$

Next, we use Example D.2 to show that the universality guarantee established for Algorithm 3 in Theorem 4.4 is essentially tight.

Example D.2. *For any $\alpha \in \left(0, \frac{1}{2}\right]$ and any integer n such that $k := \alpha n$ is an integer between 1 and $n - 1$, let $\mathcal{M} \subseteq 2^{[n]}$ be the k-uniform matroid, i.e., $\mathcal{M} = \{X \subseteq [n] \mid |X| \leq k\}$. Let $\mathcal{D}_A \in \Delta(2^{[n]})$ be the trivial prior distribution such that all elements in $[n]$ are always active.*

Proposition D.2. *In Example D.2, the prior distribution $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, but Algorithm 3 is at most $O(\alpha^2)$-balanced for $(\mathcal{M}, \mathcal{D}_A)$. Moreover, as n goes to infinity, Algorithm 3 is arbitrarily close to being $\frac{\alpha^2}{2}$-balanced for $(\mathcal{M}, \mathcal{D}_A)$.*

Proof. First, notice that in Example D.2, $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, because the CRS that outputs a uniformly random subset of k elements in $[n]$ is $\frac{k}{n}$-balanced, and $\frac{k}{n} = \alpha$.

Moreover, because of the symmetry of instance $(\mathcal{M}, \mathcal{D}_A)$, we can assume w.l.o.g. that the order preselected by Algorithm 3 is the natural order $1, 2, \ldots, n$. We denote $T_i := T \cap [i]$ for all $i \in [n]$ and $T_0 := \emptyset$, where T is the set sampled in Algorithm 3. Then, Algorithm 3 essentially enumerates all elements according to the natural order and selects each element i iff $i \in T_i$ and

$|T_{i-1}| < k$. In particular, this implies that the probability that element n is selected by Algorithm 3 is $\Pr[n \in T_n, |T_{n-1}| < k]$, and we derive that

$$\Pr[n \in T_n, |T_{n-1}| < k]$$

$$= \sum_{\ell=0}^{k-1} \Pr[n \in T_n, |T_{n-1}| = \ell]$$

$$= \sum_{\ell=0}^{k-1} \Pr[n \in T_n \mid |T_{n-1}| = \ell] \times \Pr[|T_{n-1}| = \ell]$$

$$= \sum_{\ell=0}^{k-1} \frac{\ell+1}{n+1} \times \Pr[|T_{n-1}| = \ell] \qquad \text{(By Eq. (11) in Lemma 4.3)}$$

$$= \sum_{\ell=0}^{k-1} \frac{\ell+1}{n+1} \times \Pr_{\sigma \sim \mathcal{P}([n])}[|\mathsf{prefix}(\sigma, n)| = \ell] \qquad \text{(By Lemma 4.3)}$$

$$= \sum_{\ell=0}^{k-1} \frac{\ell+1}{n+1} \times \frac{1}{n} = \frac{k(k+1)}{2n(n+1)} = \frac{\alpha(\alpha+1/n)}{2(1+1/n)},$$

where the third-to-last equality is because the probability that there are ℓ elements before element n in a uniformly random permutation of $[n]$ is $\frac{1}{n}$ for any $\ell \in \{0, 1, \ldots, n-1\}$. Thus, the probability that element n is selected by Algorithm 3 is $\frac{\alpha(\alpha+1/n)}{2(1+1/n)} = O(\alpha^2)$. This implies that Algorithm 3 is $O(\alpha^2)$-balanced for $(\mathcal{M}, \mathcal{D}_A)$. Finally, we remark that as n goes to infinity, $\frac{\alpha(\alpha+1/n)}{2(1+1/n)}$ approaches $\frac{\alpha^2}{2}$, and hence, Algorithm 3 is arbitrarily close to being $\frac{\alpha^2}{2}$-balanced for $(\mathcal{M}, \mathcal{D}_A)$. $\qquad \square$

On a separate note, the examples in this section are constructed only to demonstrate that the universality guarantees which we established for Algorithm 1 and Algorithm 3 are tight. There are simple fixes that can be made to these two algorithms to address those examples. However, it seems to us that to strengthen Algorithm 1 and Algorithm 3 to achieve $(\alpha, \Omega(\alpha))$-universality in general, we need to leverage the fact that some elements can be subsampled with a significantly higher probability than others, while keeping the probability of being spanned by the subsample sufficiently low for every element. For instance, in Example D.1, if we subsample edges in all hats with probability $\frac{\alpha}{2}$ and subsample edge (u, u') with probability 1, the probability that an edge in a hat is spanned by the subsample remains small.

E   Details of approximately solving the LPs

In this section, we discuss the details of approximately solving the LPs in Section 5 and Section 6. First, we show that the coefficients in those LPs can be estimated with sufficiently high accuracy.

Lemma E.1. *For any matroid $\mathcal{M} \subseteq 2^{[n]}$ and any prior distribution $\mathcal{D}_A \in \Delta(2^{[n]})$, given any OCRS $\mathcal{D}_\phi$, for any $\delta, \eta > 0$, we can compute random estimates $\tilde{x}_i, \tilde{q}_i \in \mathbb{Q}$ of $x_i := \Pr_{A \sim \mathcal{D}_A}[i \in A]$ and $q_i := \Pr_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi}[i \in \phi(A)]$ for all $i \in [n]$ such that*

$$\Pr[|\tilde{x}_i - x_i| \geq \eta x_i] \leq \delta \text{ and } \Pr[|\tilde{q}_i - q_i| \geq \eta x_i] \leq \delta.$$

Moreover, for each $i \in [n]$, estimates $\tilde{x}_i$ and $\tilde{q}_i$ have encoding lengths $O\left(\log\left(\frac{\log(1/\delta)}{\eta \cdot p_{\min}}\right)\right)$ and can be computed in $O\left(\frac{\log(1/\delta)}{\eta^2 p_{\min}^2} \cdot (t_{\mathcal{D}_A} + t_{\mathcal{D}_\phi})\right)$ time, where $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$, and $t_{\mathcal{D}_A}$ is the time it takes to generate a sample of $\mathcal{D}_A$, and $t_{\mathcal{D}_\phi}$ is the runtime of OCRS $\mathcal{D}_\phi$ on $(\mathcal{M}, \mathcal{D}_A)$.

Proof. For any $\eta, \delta > 0$, we sample $m = \left\lceil \frac{2\ln(2/\delta)}{\eta^2 p_{\min}^2} \right\rceil$ sets $A_1, \ldots, A_m$ from $\mathcal{D}_A$ independently, and we run OCRS $\mathcal{D}_\phi$ for each sample, which outputs sets $B_1, \ldots, B_m$. For all $i \in [n]$ and $j \in [m]$, we let $a_{i,j} = \mathbb{1}(i \in A_j)$ and $b_{i,j} = \mathbb{1}(i \in B_j)$. It is clear that $\mathbb{E}[a_{i,j}] = x_i$ and $\mathbb{E}[b_{i,j}] = q_i$ for all $i \in [n]$ and $j \in [m]$. Therefore, by Lemma A.2-ii. and our choice of m, we have that for all $i \in [n]$,

$$\Pr\left[\left|\sum_{j=1}^{m} a_{i,j} - mx_i\right| \geq \eta \cdot mx_i\right] \leq 2\exp\left(-\frac{\eta^2 mx_i^2}{2}\right) \leq 2\exp\left(-\frac{\eta^2 mp_{\min}^2}{2}\right) \leq \delta,$$

$$\Pr\left[\left|\sum_{j=1}^{m} b_{i,j} - mq_i\right| \geq \eta \cdot mx_i\right] \leq 2\exp\left(-\frac{\eta^2 mx_i^2}{2}\right) \leq 2\exp\left(-\frac{\eta^2 mp_{\min}^2}{2}\right) \leq \delta.$$

We let $\tilde{x}_i = \frac{\sum_{j=1}^{m} a_{i,j}}{m}$ and $\tilde{q}_i = \frac{\sum_{j=1}^{m} b_{i,j}}{m}$ for all $i \in [n]$. It follows from the above inequalities that $\Pr\left[|\tilde{x}_i - x_i| \geq \eta x_i\right] \leq \delta$ and $\Pr\left[|\tilde{q}_i - q_i| \geq \eta x_i\right] \leq \delta$. For each $i \in [n]$, the runtime to compute the estimates $\tilde{x}_i, \tilde{q}_i$ is $O\left(\frac{\log(1/\delta)}{\eta^2 p_{\min}^2} \cdot (t_{\mathcal{D}_A} + t_{\mathcal{D}_\phi})\right)$, because we need to generate m samples from $\mathcal{D}_A$ and run OCRS $\mathcal{D}_\phi$ for each sample. The encoding lengths of $\tilde{x}_i, \tilde{q}_i$ are $O(\log(m)) = O\left(\log\left(\frac{\log(1/\delta)}{\eta \cdot p_{\min}}\right)\right)$. $\square$

Next, we discuss how to approximately solve the LPs from Section 5 and Section 6 separately. We assume that the reader has already read those two sections.

E.1 Approximately solving the LPs in Section 5

For any $\alpha \in (0, 1]$ and $\varepsilon \in \left(0, \frac{1}{6}\right)$, given any matroid $\mathcal{M} \subseteq 2^{[n]}$ and α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, we show how to compute a $((1 - 6\varepsilon)\alpha)$-balanced OCRS with preselected order for $(\mathcal{M}, \mathcal{D}_A)$ in $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}})\right)$ time, where $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$, and $t_{\mathcal{D}_A}$ and $t_{\mathcal{M}}$ are the time it takes to generate a sample of $\mathcal{D}_A$ and to check whether a set of elements is in $\mathcal{M}$ respectively. This is achieved by solving the approximate versions of the LPs in Section 5 with estimated coefficients, which we now elaborate.

First, for each permutation $\pi \in \mathcal{S}_n$, we apply Lemma E.1 to OCRS ϕ_π (Algorithm 4) by setting $\eta = \varepsilon\alpha$ and $\delta = \frac{\varepsilon}{n(n!+1)}$, which generates random estimates $\tilde{x}_i$ and $\tilde{q}_{i,\pi}$ for all $i \in [n]$, with encoding lengths $O\left(\log\left(\frac{n}{\alpha \cdot \varepsilon \cdot p_{\min}}\right)\right)$, in time $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}})\right)$ (because the runtime of OCRS ϕ_π is $O(n \cdot t_{\mathcal{M}})$), such that

$$\Pr\left[|\tilde{x}_i - x_i| \geq \varepsilon\alpha x_i\right] \leq \frac{\varepsilon}{n(n!+1)} \text{ and } \Pr\left[|\tilde{q}_{i,\pi} - q_{i,\pi}| \geq \varepsilon\alpha x_i\right] \leq \frac{\varepsilon}{n(n!+1)}.$$

We let E denote the event that $|\tilde{x}_i - x_i| \leq \varepsilon\alpha x_i$ and $|\tilde{q}_{i,\pi} - q_{i,\pi}| \leq \varepsilon\alpha x_i$ for all $i \in [n]$ and $\pi \in \mathcal{S}_n$. By a union bound, we have that $\Pr[E] \geq 1 - \varepsilon$. (We emphasize that we generate random estimates $\tilde{x}_i$ and $\tilde{q}_{i,\pi}$ for all $i \in [n]$ and $\pi \in \mathcal{S}_n$ only for the analysis, and the algorithm will only need a polynomial number of them.)

Now we formulate an approximate version of (DP) in Eq. (17) using estimated coefficients:

$$\min_{\gamma, \mu_i} \gamma$$
$$\text{s.t.} \sum_{i \in [n]} \tilde{q}_{i,\pi}\mu_i \leq \gamma \qquad \forall \pi \in \mathcal{S}_n$$
$$\sum_{i \in [n]} \tilde{x}_i\mu_i = 1$$

34

$$\mu_i \geq 0 \qquad\qquad \forall i \in [n]. \qquad\qquad (31)$$

In the following, we establish an approximate version of Lemma 5.3.

Lemma E.2. *For any $\alpha \in (0,1]$ and $\varepsilon \in \left(0, \frac{1}{6}\right)$, if the prior distribution $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, then conditioned on event E, any vector $\mu \in \mathbb{R}^n$ that satisfies the last two constraints of the LP in Eq. (31) must also satisfy that $\sum_{i \in [n]} \tilde{q}_{i,\pi_\mu} \mu_i \geq (1 - 2\varepsilon)\alpha$, where permutation π_μ is defined in Definition 5.1.*

Proof. For any $\mu \in \mathbb{R}^n$ that satisfies the last two constraints of the LP in Eq. (31), we derive that

$$
\begin{aligned}
\sum_{i \in [n]} \tilde{q}_{i,\pi_\mu} \mu_i &\geq \sum_{i \in [n]} (q_{i,\pi_\mu} - \varepsilon\alpha x_i)\mu_i && \text{(By event } E\text{)} \\
&\geq \alpha - \sum_{i \in [n]} \varepsilon\alpha x_i \mu_i && \text{(By Lemma 5.3)} \\
&\geq \alpha - \sum_{i \in [n]} \varepsilon\alpha \cdot \frac{\tilde{x}_i \mu_i}{1 - \varepsilon\alpha} && \text{(By event } E\text{)} \\
&= \frac{1 - \varepsilon\alpha - \varepsilon}{1 - \varepsilon\alpha} \cdot \alpha && \text{(By the second constraint in the LP)} \\
&\geq (1 - \varepsilon\alpha - \varepsilon)\alpha \geq (1 - 2\varepsilon)\alpha && \text{(Since } \alpha \leq 1 \text{ and } \varepsilon \in \left(0, \frac{1}{6}\right)\text{).}
\end{aligned}
$$

$\square$

Next, we use the ellipsoid method to reduce the number of constraints in the LP in Eq. (31) such that its optimal value remains at least $(1 - 3\varepsilon)\alpha$. We consider the following polytope P_ε:

$$P_\varepsilon := \{\mu \in \mathbb{R}^n_{\geq 0} \mid \sum_{i \in [n]} \tilde{x}_i \mu_i = 1, \ \sum_{i \in [n]} \tilde{q}_{i,\pi} \mu_i \leq (1 - 3\varepsilon)\alpha \text{ for all } \pi \in \mathcal{S}_n\}.$$

Since we assumed that $\alpha \in (0,1]$ and $\varepsilon \in \left(0, \frac{1}{6}\right)$, by Lemma E.2, conditioned on event E, any vector $\mu \in \mathbb{R}^n_{\geq 0}$ such that $\sum_{i \in [n]} \tilde{x}_i \mu_i = 1$ must satisfy that $\sum_{i \in [n]} \tilde{q}_{i,\pi_\mu} \mu_i \geq (1 - 2\varepsilon)\alpha$. Therefore, conditioned on event E, polytope P_ε is empty, and we can construct an efficient separation oracle for P_ε as follows: Given any $\mu \in \mathbb{R}^n_{\geq 0}$ such that $\sum_{i \in [n]} \tilde{x}_i \mu_i = 1$, the oracle outputs the violated constraint $\sum_{i \in [n]} \tilde{q}_{i,\pi_\mu} \mu_i \leq (1 - 3\varepsilon)\alpha$. Using this separation oracle, we can apply the ellipsoid method to identify a subset of permutations $\mathcal{S}'_n \subseteq \mathcal{S}_n$ such that the following polytope P'_ε is empty:

$$P'_\varepsilon = \{\mu \in \mathbb{R}^n_{\geq 0} \mid \sum_{i \in [n]} \tilde{x}_i \mu_i = 1, \ \sum_{i \in [n]} \tilde{q}_{i,\pi} \mu_i \leq (1 - 3\varepsilon)\alpha \text{ for all } \pi \in \mathcal{S}'_n\}.$$

In particular, when we apply the ellipsoid method, we can compute the estimates $\tilde{x}_i$ for all $i \in [n]$ initially and compute the estimates $\tilde{q}_{i,\pi}$ only if necessary. That is, for any $\pi \in \mathcal{S}_n$, we compute estimates $\tilde{q}_{i,\pi}$ for all $i \in [n]$ only if the separation oracle needs to output the constraint $\sum_{i \in [n]} \tilde{q}_{i,\pi} \mu_i \leq (1 - 3\varepsilon)\alpha$. Since the total number of iterations of the ellipsoid method is polynomial in the number of variables and the maximum encoding length of the coefficients in the polytope constraints (Grötschel et al., 2012), it takes $O\left(\mathsf{poly}\left(n \log\left(\frac{n}{\alpha \cdot \varepsilon \cdot p_{\min}}\right)\right)\right)$ iterations to certify that polytope P_ε is empty. It follows that $|\mathcal{S}'_n| = O\left(\mathsf{poly}\left(n \log\left(\frac{n}{\alpha \cdot \varepsilon \cdot p_{\min}}\right)\right)\right)$. The overall runtime of the ellipsoid method, accounting for the time to compute the estimates, is $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathcal{D}_A} + t_{\mathcal{M}})\right)$.

Finally, we consider the following LP (LP') and its dual (DP').

$$
\begin{aligned}
\text{(LP')} \quad &\max_{\beta,\,\lambda_\pi} \beta \\
&\text{s.t.} \sum_{\pi\in\mathcal{S}'_n} \tilde{q}_{i,\pi}\lambda_\pi \geq \beta\tilde{x}_i \qquad \forall i\in[n] \\
&\qquad \sum_{\pi\in\mathcal{S}'_n} \lambda_\pi = 1 \\
&\qquad \lambda_\pi \geq 0 \qquad\qquad\qquad \forall\pi\in\mathcal{S}'_n.
\end{aligned}
$$

$$
\begin{aligned}
\text{(DP')} \quad &\min_{\gamma,\,\mu_i} \gamma \\
&\text{s.t.} \sum_{i\in[n]} \tilde{q}_{i,\pi}\mu_i \leq \gamma \qquad \forall\pi\in\mathcal{S}'_n \\
&\qquad \sum_{i\in[n]} \tilde{x}_i\mu_i = 1 \\
&\qquad \mu_i \geq 0 \qquad\qquad\qquad \forall i\in[n].
\end{aligned} \tag{32}
$$

Because polytope P'_ε is empty conditioned on event E, the optimal value of (DP') in Eq. (32) is at least $(1-3\varepsilon)\alpha$ conditioned on E. By LP duality, the optimal value of (LP') is also at least $(1-3\varepsilon)\alpha$ conditioned on E. Note that (LP') has $O\left(\mathsf{poly}\left(n\log\left(\frac{n}{\alpha\cdot\varepsilon\cdot p_{\min}}\right)\right)\right)$ variables and constraints, and hence, its optimal solution, which we denote by (β^*,λ^*_π), can be computed in $O\left(\mathsf{poly}\left(n\log\left(\frac{n}{\alpha\cdot\varepsilon\cdot p_{\min}}\right)\right)\right)$ time. By the last two constraints in (LP'), variables λ^*_π for all $\pi\in\mathcal{S}'_n$ together specify a distribution $\mathcal{D}^*_\pi$ over $\mathcal{S}'_n$. We observe that ϕ_π with $\pi\sim\mathcal{D}^*_\pi$ is a randomized OCRS with preselected order for $(\mathcal{M},\mathcal{D}_A)$. This OCRS is $((1-5\varepsilon)\alpha)$-balanced conditioned on event E, because for all $i\in[n]$, we have that

$$
\begin{aligned}
\sum_{\pi\in\mathcal{S}'_n} q_{i,\pi}\lambda^*_\pi &\geq \sum_{\pi\in\mathcal{S}'_n} (\tilde{q}_{i,\pi}-\varepsilon\alpha x_i)\lambda^*_\pi && \text{(By event } E) \\
&= \sum_{\pi\in\mathcal{S}'_n} \tilde{q}_{i,\pi}\lambda^*_\pi - \varepsilon\alpha x_i && \text{(By the second constraint in (LP'))} \\
&\geq \beta^*\tilde{x}_i - \varepsilon\alpha x_i && \text{(By the first constraint in (LP'))} \\
&\geq \beta^*(1-\varepsilon\alpha)x_i - \varepsilon\alpha x_i && \text{(By event } E) \\
&\geq (1-3\varepsilon)\cdot(1-\varepsilon\alpha)\alpha x_i - \varepsilon\alpha x_i && \text{(Since } \beta^*\geq(1-3\varepsilon)\alpha \text{ conditioned on } E) \\
&\geq (1-3\varepsilon-\varepsilon\alpha)\alpha x_i - \varepsilon\alpha x_i \\
&\geq (1-5\varepsilon)\alpha x_i.
\end{aligned}
$$

Taking into account the probability of event E, which is at least $1-\varepsilon$, this OCRS is $((1-\varepsilon)\cdot(1-5\varepsilon)\alpha)$-balanced, and we note that $(1-\varepsilon)\cdot(1-5\varepsilon)\geq 1-6\varepsilon$. The total runtime of using the ellipsoid method to find $\mathcal{S}'_n$ and solving (LP') in Eq. (32) is $O\left(\mathsf{poly}\left(\frac{n}{\alpha\cdot\varepsilon\cdot p_{\min}}\right)\cdot(t_{\mathcal{D}_A}+t_{\mathcal{M}})\right)$.

E.2 Approximately solving the LPs in Section 6

For any $\alpha,c\in(0,1]$ and $\varepsilon\in\left(0,\frac{1}{7}\right)$, given any c-competitive matroid secretary algorithm ALG in any arrival model, any matroid $\mathcal{M}\subseteq 2^{[n]}$ and α-uncontentious distribution $\mathcal{D}_A$ for $\mathcal{M}$, we show how to compute a $((1-7\varepsilon)c\alpha)$-balanced OCRS for $(\mathcal{M},\mathcal{D}_A)$ in $O\left(\mathsf{poly}\left(\frac{n}{\alpha\cdot c\cdot\varepsilon\cdot p_{\min}}\right)\cdot(t_{\mathsf{ALG}}+t_{\mathcal{D}_A})\right)$

time, where $p_{\min} := \min_{i \in [n]} \Pr_{A \sim \mathcal{D}_A}[i \in A]$, and t_{ALG} is the worst-case runtime of algorithm ALG on matroid secretary problem instances specified by matroid $\mathcal{M}$ and weight vectors $w \in W_\varepsilon^n$, and $t_{\mathcal{D}_A}$ is the time it takes to generate a sample of $\mathcal{D}_A$. This is achieved by solving the approximate versions of the LPs in Section 6 with estimated coefficients, which we now elaborate.

First, for each vector $w \in W_\varepsilon^n$, we apply Lemma E.1 to OCRS $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ by setting $\eta = \varepsilon c \alpha$ and $\delta = \frac{\varepsilon}{n(|W_\varepsilon|^n+1)}$ (recall that $|W_\varepsilon| = O\left(\frac{n}{\varepsilon \cdot p_{\min}}\right)$), which generates random estimates $\tilde{x}_i$ and $\tilde{q}_{i,w}$ for all $i \in [n]$, with encoding lengths $O\left(\log\left(\frac{n}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right)\right)$, in time $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathsf{ALG}} + t_{\mathcal{D}_A})\right)$, such that

$$\Pr\left[|\tilde{x}_i - x_i| \geq \varepsilon c \alpha x_i\right] \leq \frac{\varepsilon}{n(|W_\varepsilon|^n+1)} \text{ and } \Pr\left[|\tilde{q}_{i,w} - q_{i,w}| \geq \varepsilon c \alpha x_i\right] \leq \frac{\varepsilon}{n(|W_\varepsilon|^n+1)}.$$

We let E denote the event that $|\tilde{x}_i - x_i| \leq \varepsilon c \alpha x_i$ and $|\tilde{q}_{i,w} - q_{i,w}| \leq \varepsilon c \alpha x_i$ for all $i \in [n]$ and $w \in W_\varepsilon^n$. By a union bound, we have that $\Pr[E] \geq 1 - \varepsilon$. (We emphasize that we generate random estimates $\tilde{x}_i$ and $\tilde{q}_{i,w}$ for all $i \in [n]$ and $w \in W_\varepsilon^n$ only for the analysis, and the algorithm will only need a polynomial number of them.)

Now we formulate an approximate version of (DP1) in Eq. (18) using estimated coefficients:

$$
\begin{aligned}
\text{(DP1')} \qquad & \min_{\gamma, \mu_i} \gamma \\
& \text{s.t.} \sum_{i \in [n]} \tilde{q}_{i,w} \mu_i \leq \gamma \qquad && \forall w \in W_\varepsilon^n \\
& \qquad \sum_{i \in [n]} \tilde{x}_i \mu_i = 1 \\
& \qquad \mu_i \geq 0 \qquad && \forall i \in [n].
\end{aligned}
\tag{33}
$$

In the following, we establish an approximate version of Lemma 6.2.

Lemma E.3. *For any $\alpha, c \in (0,1]$ and $\varepsilon \in \left(0, \frac{1}{7}\right)$, if the matroid secretary algorithm ALG is c-competitive, and the prior distribution $\mathcal{D}_A$ is α-uncontentious for matroid $\mathcal{M}$, then conditioned on event E, any vector $\mu \in \mathbb{R}^n$ that satisfies the last two constraints in (DP1') must also satisfy that $\sum_{i \in [n]} \tilde{q}_{i,\mu'} \mu_i \geq (1 - 3\varepsilon) c \alpha$, where $\mu' \in W_\varepsilon^n$ is defined in Eq. (19).*

Proof. For any $\mu \in \mathbb{R}^n$ that satisfies the last two constraints of (DP1') in Eq. (33), we let $\mu' \in W_\varepsilon^n$ be the corresponding vector defined in Eq. (19), and we derive that

$$
\begin{aligned}
\sum_{i \in [n]} \tilde{q}_{i,\mu'} \mu_i &\geq \sum_{i \in [n]} (q_{i,\mu'} - \varepsilon c \alpha x_i) \mu_i && \text{(By event } E) \\
&\geq (1 - \varepsilon) c \alpha - \sum_{i \in [n]} \varepsilon c \alpha x_i \mu_i && \text{(By Lemma 6.2)} \\
&\geq (1 - \varepsilon) c \alpha - \sum_{i \in [n]} \varepsilon c \alpha \cdot \frac{\tilde{x}_i \mu_i}{1 - \varepsilon c \alpha} && \text{(By event } E) \\
&= \frac{(1 - \varepsilon)(1 - \varepsilon c \alpha) - \varepsilon}{1 - \varepsilon c \alpha} \cdot c \alpha && \text{(By the second constraint in the (DP1'))} \\
&\geq ((1 - \varepsilon)(1 - \varepsilon c \alpha) - \varepsilon) c \alpha && \text{(Since } \alpha, c \leq 1 \text{ and } \varepsilon \in \left(0, \frac{1}{7}\right)) \\
&\geq ((1 - \varepsilon - \varepsilon c \alpha) - \varepsilon) c \alpha \geq (1 - 3\varepsilon) c \alpha.
\end{aligned}
$$

$\square$

Next, we use the ellipsoid method to reduce the number of constraints in (DP1') such that its optimal value remains at least $(1 - 4\varepsilon)c\alpha$. We consider the following polytope $\tilde{Q}_\varepsilon$:

$$\tilde{Q}_\varepsilon := \{\mu \in \mathbb{R}_{\geq 0}^n \mid \sum_{i \in [n]} \tilde{x}_i \mu_i = 1, \sum_{i \in [n]} \tilde{q}_{i,w} \mu_i \leq (1 - 4\varepsilon)c\alpha \text{ for all } w \in W_\varepsilon^n\}.$$

Since we assumed that $\alpha, c \in (0, 1]$ and $\varepsilon \in \left(0, \frac{1}{7}\right)$, by Lemma E.3, conditioned on event E, any vector $\mu \in \mathbb{R}_{\geq 0}^n$ such that $\sum_{i \in [n]} \tilde{x}_i \mu_i = 1$ must satisfy that $\sum_{i \in [n]} \tilde{q}_{i,\mu'} \mu_i \geq (1 - 3\varepsilon)c\alpha$, where μ' is defined in Eq. (19). Therefore, conditioned on event E, polytope $\tilde{Q}_\varepsilon$ is empty, and we can construct an efficient separation oracle for $\tilde{Q}_\varepsilon$ as follows: Given any $\mu \in \mathbb{R}_{\geq 0}^n$ such that $\sum_{i \in [n]} \tilde{x}_i \mu_i = 1$, the oracle outputs the violated constraint $\sum_{i \in [n]} \tilde{q}_{i,\mu'} \mu_i \leq (1 - 4\varepsilon)c\alpha$ for μ' given by Eq. (19). Using this separation oracle, we can apply the ellipsoid method to identify a subset of vectors $\tilde{W} \subseteq W_\varepsilon^n$ such that the following polytope $\tilde{Q}_\varepsilon'$ is empty:

$$\tilde{Q}_\varepsilon' = \{\mu \in \mathbb{R}_{\geq 0}^n \mid \sum_{i \in [n]} \tilde{x}_i \mu_i = 1, \sum_{i \in [n]} \tilde{q}_{i,w} \mu_i \leq (1 - 4\varepsilon)c\alpha \text{ for all } w \in \tilde{W}\}.$$

In particular, when we apply the ellipsoid method, we can compute the estimates $\tilde{x}_i$ for all $i \in [n]$ initially and compute the estimates $\tilde{q}_{i,w}$ only if necessary. That is, for any $w \in W_\varepsilon^n$, we compute estimates $\tilde{q}_{i,w}$ for all $i \in [n]$ only if the separation oracle needs to output the constraint $\sum_{i \in [n]} \tilde{q}_{i,w} \mu_i \leq (1 - 4\varepsilon)c\alpha$. Since the total number of iterations of the ellipsoid method is polynomial in the number of variables and the maximum encoding length of the coefficients in the polytope constraints (Grötschel et al., 2012), it takes $O\left(\mathsf{poly}\left(n \log \left(\frac{1}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right)\right)\right)$ iterations to certify that polytope $\tilde{Q}_\varepsilon$ is empty. It follows that $|\tilde{W}| = O\left(\mathsf{poly}\left(n \log \left(\frac{1}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right)\right)\right)$. The overall runtime of the ellipsoid method, accounting for the time to compute the estimates, is $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathsf{ALG}} + t_{\mathcal{D}_A})\right)$.

Finally, we consider the following LP (LP2') and its dual (DP2').

$$\begin{aligned}
\text{(LP2')} \quad & \max_{\beta, \lambda_w} \beta \\
& \text{s.t.} \sum_{w \in \tilde{W}} \tilde{q}_{i,w} \lambda_w \geq \beta \tilde{x}_i \quad && \forall i \in [n] \\
& \qquad \sum_{w \in \tilde{W}} \lambda_w = 1 \\
& \qquad \lambda_w \geq 0 && \forall w \in \tilde{W}.
\end{aligned}$$

$$\begin{aligned}
\text{(DP2')} \quad & \min_{\gamma, \mu_i} \gamma \\
& \text{s.t.} \sum_{i \in [n]} \tilde{q}_{i,w} \mu_i \leq \gamma \quad && \forall w \in \tilde{W} \\
& \qquad \sum_{i \in [n]} \tilde{x}_i \mu_i = 1 \\
& \qquad \mu_i \geq 0 && \forall i \in [n].
\end{aligned}$$

Because polytope $\tilde{Q}_\varepsilon'$ is empty conditioned on event E, the optimal value of (DP2') is at least $(1 - 4\varepsilon)c\alpha$ conditioned on E. By LP duality, the optimal value of (LP2') is also at least $(1 - 4\varepsilon)c\alpha$ conditioned on E. Since (LP2') has $O\left(\mathsf{poly}\left(n \log \left(\frac{1}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right)\right)\right)$ variables and constraints, its optimal solution, which we denote by (β^*, λ_w^*), can be computed in $O\left(\mathsf{poly}\left(n \log \left(\frac{1}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right)\right)\right)$ time.

By the last two constraints in (LP2'), variables λ_w^* for all $w \in \tilde{W}$ together specify a distribution $\mathcal{D}_w^*$ over $\tilde{W}$. We observe that $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ with $w \sim \mathcal{D}_w^*$ is an OCRS for $(\mathcal{M}, \mathcal{D}_A)$ in the same arrival model as algorithm ALG. This OCRS is $((1-6\varepsilon)c\alpha)$-balanced conditioned on event E, because for all $i \in [n]$, we have that

$$
\begin{aligned}
\sum_{w \in \tilde{W}} q_{i,w}\lambda_w^* &\geq \sum_{w \in \tilde{W}} (\tilde{q}_{i,w} - \varepsilon c\alpha x_i)\lambda_w^* && \text{(By event } E\text{)} \\
&= \sum_{w \in \tilde{W}} \tilde{q}_{i,w}\lambda_w^* - \varepsilon c\alpha x_i && \text{(By the second constraint in (LP2'))} \\
&\geq \beta^* \tilde{x}_i - \varepsilon c\alpha x_i && \text{(By the first constraint in (LP2'))} \\
&\geq \beta^*(1 - \varepsilon c\alpha)x_i - \varepsilon c\alpha x_i && \text{(By event } E\text{)} \\
&\geq (1-4\varepsilon) \cdot (1 - \varepsilon c\alpha)c\alpha x_i - \varepsilon c\alpha x_i && \text{(Since } \beta^* \geq (1-4\varepsilon)c\alpha \text{ conditioned on } E\text{)} \\
&\geq (1-4\varepsilon - \varepsilon c\alpha)c\alpha x_i - \varepsilon c\alpha x_i \\
&\geq (1-5\varepsilon)c\alpha x_i - \varepsilon c\alpha x_i = (1-6\varepsilon)c\alpha x_i.
\end{aligned}
$$

Taking into account the probability of event E, which is at least $1 - \varepsilon$, this OCRS is $((1-\varepsilon) \cdot (1-6\varepsilon)c\alpha)$-balanced, and we note that $(1-\varepsilon) \cdot (1-6\varepsilon) \geq 1 - 7\varepsilon$. The total runtime of using the ellipsoid method to find $\tilde{W}$ and solving (LP2') is $O\left(\mathsf{poly}\left(\frac{n}{\alpha \cdot c \cdot \varepsilon \cdot p_{\min}}\right) \cdot (t_{\mathsf{ALG}} + t_{\mathcal{D}_A})\right)$.

Remark E.1. *In fact, what we (and Dughmi (2020)) have established above is a reduction from universal online contention resolution to a simpler type of matroid secretary problem. Specifically, note that (implicitly) in the definition of $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ in Section 6, the matroid secretary algorithm ALG is aware of the weight vector w and the prior distribution $\mathcal{D}_A$ from the outset, and it can use these additional information to make decisions. What we have proved above is that $\mathcal{D}_\phi^{(\mathsf{ALG},w)}$ with $w \sim \mathcal{D}_w^*$ is $((1-7\varepsilon)c\alpha)$-balanced for $(\mathcal{M}, \mathcal{D}_A)$, as long as ALG satisfies the inequality stated in Lemma 6.1, i.e., for all $w \in \mathbb{R}_{\geq 0}^n$,*

$$
\mathbb{E}_{A \sim \mathcal{D}_A, \phi \sim \mathcal{D}_\phi^{(\mathsf{ALG},w)}}\left[\sum\nolimits_{i \in \phi(A)} w_i\right] \geq c \cdot \alpha \cdot \mathbb{E}_{A \sim \mathcal{D}_A}\left[\sum\nolimits_{i \in A} w_i\right].
$$

From this perspective, Theorem 5.4 can be viewed as a corollary of the above reduction, because Lemma 5.2 essentially shows that the above inequality holds with $c = 1$, if ALG greedily selects active elements in the decreasing order of their weights.