Secretary Problems with Convex Costs

Siddharth Barman Seeun Umboh Shuchi Chawla
David Malec

University of Wisconsin–Madison
{sid, seeun, shuchi, dmalec}@cs.wisc.edu

Abstract

We consider online resource allocation problems where given a set of requests our goal is to select a subset that maximizes a value minus cost type of objective function. Requests are presented online in random order, and each request possesses an adversarial value and an adversarial size. The online algorithm must make an irrevocable accept/reject decision as soon as it sees each request. The "profit" of a set of accepted requests is its total value minus a convex cost function of its total size. This problem generalizes the so-called knapsack secretary problem and is closely related to the submodular secretary problem. Unlike previous work on secretary problems, one of the main challenges we face is that the objective function can be positive or negative and we must guard against accepting requests that look good early on but cause the solution to have an arbitrarily large cost as more requests are accepted.

We study this problem under various feasibility constraints and present online algorithms with competitive ratios only a constant factor worse than those known in the absence of costs for the same feasibility constraints. We also consider a multi-dimensional version of the problem that generalizes multi-dimensional knapsack within a secretary framework. In the absence of any feasibility constraints, we present an $O(\ell)$ competitive algorithm where ℓ is the number of dimensions; this matches within constant factors the best known ratio for multi-dimensional knapsack secretary.

1 Introduction

Numerous real-world resource allocation problems require online decision making. Consider, for example, a wireless access point accepting connections from mobile nodes, or a cloud computing service accepting jobs; the server in either case must accept or reject these requests and allocate resources appropriately without knowing the quality of future requests. In this work we focus on online problems where the server makes binary decisions of accepting or rejecting requests as soon as they are received, and these decisions are irrevocable. Other examples which require online decision making include sponsored search auctions where advertisers must decide which keywords to bid on without knowing about future search queries, and sensor networks where representative samples have to be selected from an online heterogeneous stream.

A classical example of online decision making is the secretary problem. Here a company is interested in hiring a candidate for a single position; candidates arrive for interview in *random order*, and the company must accept or reject each candidate following the interview. The goal is to select the best candidate as often as possible. What makes the problem challenging is that each interview merely reveals the rank of the candidate relative to the ones seen previously, but not the ones following. Nevertheless, Dynkin [10] showed that it is possible to succeed with constant probability using the following algorithm: unconditionally reject the first $1/e$ fraction of the candidates; then hire the next candidate that is better than all of the ones seen previously. Dynkin showed that as the number of candidates goes to infinity, this algorithm hires the best candidate with probability approaching $1/e$ and in fact this is the best possible.

More general resource allocation settings may allow picking multiple candidates subject to a certain feasibility constraint. We call such a problem a generalized secretary problem (GSP) and use $(\Phi, \mathcal{F})$ to denote an instance of the problem. Here $\mathcal{F}$ denotes a feasibility constraint that the set of accepted requests must satisfy (e.g. the size of the set cannot exceed a given bound), and Φ denotes an objective function that we wish to maximize. As in the classical setting, we assume that requests arrive in random order; the feasibility constraint $\mathcal{F}$ is known in advance but the quality of each request, in particular its contribution to Φ, is only revealed when the request arrives. Our goal is to minimize the competitive ratio of the online algorithm—the ratio of the objective function value at the overall optimal solution to the one obtained by the online algorithm.

Recent work has explored variants of the GSP where Φ is the sum over the accepted requests of the "value" of each request. For such a sum-of-values objective, constant factor competitive ratios are known for various kinds of feasibility constraints including cardinality constraints [16], knapsack constraints [4], and certain matroid constraints [5].

In applications such as those mentioned above, the linear sum-of-values objective does not adequately capture the tradeoffs that the algorithm faces in accepting or rejecting a request, and feasibility constraints provide only a rough approximation. Consider, for example, a wireless access point accepting connections. Each accepted request improves resource utilization and brings value to the access point. However as the number of accepted requests grows the access point performs greater multiplexing of the spectrum, and must use more and more transmitting power in order to maintain a reasonable connection bandwidth for each request. The power consumption and its associated cost are *non-linear* functions of the total load on the access point. This directly translates into a value minus cost type of objective function where the cost is an increasing function of the load or total size of all the requests accepted. We assume that the cost function is a convex function[1].

[1]Convexity is crucial in obtaining any non-trivial competitive ratio—if the cost function were concave, the only

Our goal then is to accept a set A out of a universe U of requests such that the "profit" $\pi(A) = v(A) - \mathsf{C}(s(A))$ is maximized; here $v(A)$ is the total value of all requests in A, $s(A)$ is the total size, and C is a known increasing convex cost function. Note that the problem $(\pi, 2^U)$ (i.e. where the feasibility constraint is trivial) is a generalization of the knapsack secretary problem [4] where the goal is to maximize a sum-of-values objective subject to a knapsack constraint. We further consider objectives that generalize the ℓ-dimensional knapsack secretary problem. Here, we are given ℓ different (known) convex cost functions C_i for $1 \leq i \leq \ell$, and each request is endowed with ℓ sizes, one for each dimension. The profit of a set is given by $\pi(A) = v(A) - \sum_{i=1}^{\ell} \mathsf{C}_i(s_i(A))$ where $s_i(A)$ is the total size of the set in dimension i.

We consider the profit maximization problem under various feasibility constraints. For single-dimensional costs, we obtain online algorithms with competitive ratios within a constant factor of those achievable for a sum-of-values objective with the same feasibility constraints. For ℓ-dimensional costs, in the absence of any constraints, we obtain an $O(\ell)$ competitive ratio. We remark that this is essentially the best approximation achievable even in the offline setting: Dean et al. [8] show an $\Omega(\ell^{1-\epsilon})$ hardness for the simpler ℓ-dimensional knapsack problem under a standard complexity-theoretic assumption. For the multi-dimensional problem with general feasibility constraints, our competitive ratios are worse by a factor of $O(\ell^5)$ over the corresponding versions without costs. Improving this factor is a possible avenue for future research.

We remark that the profit function π is a submodular function. Recently Bateni et al. [6] and Gupta et al. [14] studied versions of the GSP where the objective Φ is a non-monotone submodular function and provided competitive algorithms for different kinds of feasibility constraints. However, both these works make the crucial assumption that the objective is always nonnegative; it therefore does not capture π as a special case. In particular, if Φ is a monotone increasing submodular function (that is, if adding more elements to the solution cannot decrease its objective value), then to obtain a good competitive ratio it suffices to show that the online solution captures a good fraction of the optimal solution. In the case of [6] and [14], the objective function is not necessarily monotone. Nevertheless, nonnegativity implies that the universe of elements can be divided into two parts, over each of which the objective essentially behaves like a monotone submodular function in the sense that adding extra elements to a good subset of the optimal solution does not decrease its objective function value. In our setting, in contrast, adding elements with too large a size to the solution can cause the cost of the solution to become too large and therefore imply a negative profit, even if the rest of the elements are good in terms of their value-size tradeoff. As a consequence we can only guarantee good profit when no "bad" elements are added to the solution, and must ensure that this holds with high probability. This necessitates designing new techniques.

Our techniques. At a very high level our approach follows Dynkin's algorithm—we use the first few elements in the randomly ordered stream of requests to determine a threshold for classifying elements as "good" or "bad"; good elements are those with a high value to size ratio (a.k.a. density). We then apply this threshold to the remaining stream and accept all the elements that cross the threshold (and are therefore good). In the absence of any feasibility constraints (see Section 3) we show that there always exists a threshold such that the elements crossing it are a large fraction of the overall optimal solution and therefore give a good approximation. We further show how to determine this density threshold from the sample with high probability. With general feasibility constraints, it is no longer true that an arbitrary feasible subset of the good elements

solutions with a nonnegative objective function value may be to accept everything or nothing.

is a good approximation. Instead, we decompose the profit function into two parts each of which can be optimized by maximizing a certain sum-of-values function (see Section 4). We then employ previous work on GSP with a sum-of-values function to obtain a good approximation to one or the other component of profit. The decomposition of profit once again depends crucially on finding a good density threshold for classifying elements as good and bad; the two sum-of-values objectives are nonnegative only on the good elements, and we therefore filter out bad elements as a first step before applying online algorithms for those objectives.

We note that while the exposition in Section 4 focuses on a matroid feasibility constraint, the results of that section extend to any downwards-closed feasibility constraint that admits good offline and online algorithms with a sum-of-values objective[2].

In the multi-dimensional setting (discussed in Section 5), elements have different sizes along different dimensions. Therefore, a single density does not capture the value-size tradeoff that an element offers. Instead we can decompose the value of an element into ℓ different values, one for each dimension, and define densities in each dimension accordingly. This decomposes the profit across dimensions as well. Then, at a loss of a factor of ℓ, we can approximate the profit objective along the "best" dimension. The problem with this approach is that a solution that is good (or even best) in one dimension may in fact be terrible with respect to the overall profit, if its profit along other dimensions is negative. Surprisingly we show that it is possible to partition values across dimensions in such a way that there is a *single* ordering over elements in terms of their value-size tradeoff that is respected in each dimension; this allows us to prove that a solution that is good in one dimension is also good in other dimensions. We present an $O(\ell)$ competitive algorithm for the unconstrained setting based on this approach in Section 5, and defer a discussion of the constrained setting to Appendix E.

Related work. The classical secretary problem has been studied extensively; see [12, 13] and [20] for a survey. Recently a number of papers have explored variants of the GSP with a sum-of-values objective. Kleinberg [16] considered the variant where up to k secretaries can be selected (a.k.a. the k-secretary problem) and gave a $1 - O(1/\sqrt{k})$ competitive algorithm. Babaioff et al. [4] generalized this to a setting where different candidates have different sizes and the total size of the selected set must be bounded by a given amount, and gave a constant factor approximation. In [5] Babaioff et al. considered another generalization of the k-secretary problem to matroid feasibility constraints. A matroid is a set system over U that is downwards closed (that is, subsets of feasible sets are feasible), and satisfies a certain exchange property (see [18] for a comprehensive treatment). They presented an $O(\log r)$ competitive algorithm, where r is the rank of the matroid, or the size of a maximal feasible set. Subsequently, several papers improved upon the competitive ratio for special classes of matroids [1, 9, 17]. Bateni et al. [6] and Gupta et al. [14] were the first to (independently) consider non-linear objectives in this context. They gave online algorithms for non-monotone nonnegative submodular objective functions with competitive ratios within constant factors of the ratios known for the sum-of-values objective under the same feasibility constraint. Other versions of the problem that have been studied recently include: settings where elements are drawn from known or unknown distributions but arrive in an adversarial order [7, 15, 19], versions where values are permuted randomly across elements of a non-symmetric set system [21], and settings where the algorithm is allowed to reverse some of its decisions at a cost [2, 3].

[2]We obtain an $O(\alpha^4\beta)$ competitive algorithm where α is the best offline approximation and β is the best online competitive ratio for the sum-of-values objective.

2 Notation and Preliminaries

We consider instances of the generalized secretary problem represented by the pair $(\pi, \mathcal{F})$, and an implicit number n of requests or elements that arrive in an online fashion. U denotes the universe of elements. $\mathcal{F} \subset 2^U$ is a known downwards-closed feasibility constraint. Our goal is to accept a subset of elements $A \subset U$ with $A \in \mathcal{F}$ such that the objective function $\pi(A)$ is maximized. For a given set $T \subset U$, we use $O^*(T) = \text{argmax}_{A \in \mathcal{F} \cap 2^T} \pi(A)$ to denote the optimal solution over T; O^* is used as shorthand for $O^*(U)$. We now describe the function π.

In the single-dimensional cost setting, each element $e \in U$ is endowed with a value $v(e)$ and a size $s(e)$. Values and sizes are integral and are a priori unknown. The size and value functions extend to sets of elements as $s(A) = \sum_{e \in A} s(e)$ and $v(A) = \sum_{e \in A} v(e)$. Then the "profit" of a subset is given by $\pi(A) = v(A) - \mathsf{C}(s(A))$ where C is a convex function on size: $\mathsf{C} : \mathbb{Z}^+ \to \mathbb{Z}^+$. The following quantities will be useful in our analysis:

- The *density* of an element, $\rho(e) := v(e)/s(e)$. We assume without loss of generality that densities of elements are unique and denote the unique element with density γ by e_γ.
- The *marginal cost* function, $\mathsf{c}(s) := \mathsf{C}(s) - \mathsf{C}(s-1)$. Note that this is an increasing function.
- The *inverse marginal cost* function, $\bar{\mathsf{s}}(\rho)$ which is defined to be the maximum size for which an element of density ρ will have a non-negative profit increment, that is, the maximum s for which $\rho \geq \mathsf{c}(s)$.
- The *density prefix* for a given density γ and a set T, $P_\gamma^T := \{e \in T : \rho(e) \geq \gamma\}$, and the partial density prefix, $\bar{P}_\gamma^T := P_\gamma^T \setminus \{e_\gamma\}$. We use P_γ and $\bar{P}_\gamma$ as shorthand for P_γ^U and $\bar{P}_\gamma^U$ respectively.

We will sometimes find it useful to discuss fractional relaxations of the offline problem of maximizing π subject to $\mathcal{F}$. To this end, we extend the definition of subsets of U to allow for fractional membership. We use αe to denote an α-fraction of element e; this has value $v(\alpha e) = \alpha v(e)$ and size $s(\alpha e) = \alpha s(e)$. We say that a fractional subset A is feasible if its support supp(A) is feasible. Note that when the feasibility constraint can be expressed as a set of linear constraints, this relaxation is more restrictive than the linear relaxation.

Note that since costs are a convex increasing function of size, it may at times be more profitable to accept a fraction of an element rather than the whole. That is, $\text{argmax}_\alpha \pi(\alpha e)$ may be strictly less than 1. For such elements, $\rho(e) < \mathsf{c}(s(e))$. We use $\mathbb{F}$ to denote the set of all such elements: $\mathbb{F} = \{e \in U : \text{argmax}_\alpha \pi(\alpha e) < 1\}$, and $\mathbb{I} = U \setminus \mathbb{F}$ to denote the remaining elements. Our solutions will generally approximate the optimal profit from $\mathbb{F}$ by running Dynkin's algorithm for the classical secretary problem; most of our analysis will focus on $\mathbb{I}$. Let $F^*(T)$ denote the optimal (feasible) fractional subset of $T \cap \mathbb{I}$ for a given set T. Then $\pi(F^*(T)) \geq \pi(O^*(T \cap \mathbb{I}))$. We use F^* as shorthand for $F^*(U)$, and let s^* be the size of this solution.

In the multi-dimensional setting each element has an ℓ-dimensional size $s(e) = (s_1(e), \ldots, s_\ell(e))$. The cost function is composed of ℓ different convex functions, $\mathsf{C}_i : \mathbb{Z}^+ \to \mathbb{Z}^+$. The cost of a set of elements is defined to be $\mathsf{C}(A) = \sum_i \mathsf{C}_i(s_i(A))$ and the profit of A, as before, is its value minus its cost: $\pi(A) = v(A) - \mathsf{C}(A)$.

Balanced Sampling. Our algorithms learn the distribution of element values and sizes by observing the first few elements. Because of the random order of arrival, these elements form a random subset of the universe U. The following concentration result is useful in formalizing the representativeness of the sample (see Appendix A for a proof).

Lemma 1. *Given constant $c \geq 3$ and a set of elements I with associated non-negative weights, w_i for $i \in I$, say we construct a random subset J by selecting elements from I uniformly at random with probability $1/2$. If for all $k \in I$, $w_k \leq \frac{1}{c}\sum_{i \in I} w_i$ then the following inequality holds with probability at least 0.76:*

$$\sum_{j \in J} w_j \geq \beta(c) \sum_{i \in I} w_i,$$

where $\beta(c)$ is a non-decreasing function of c (and furthermore is independent of I).

3 Unconstrained Profit Maximization

We begin by developing an algorithm for the unconstrained version of the generalized secretary problem with $\mathcal{F} = 2^U$, which already exhibits some of the challenges of the general setting. Note that this setting captures as a special case the knapsack secretary problem of [4] where the goal is to maximize the total value of a subset of size at most a given bound. In fact in the offline setting, the generalized secretary problem is very similar to the knapsack problem. If all elements have the same (unit) size, then the optimal offline algorithm orders elements in decreasing order of value and picks the largest prefix in which each element contributes a positive marginal profit. When element sizes are different, a similar approach works: we order elements by density this time, and note that either a prefix of this ordering or a single element is a good approximation (much like the greedy 2-approximation for knapsack). Appendix B provides a detailed analysis of the algorithm as well as other missing proofs.

Precisely, we show that $|O^* \cap \mathbb{F}| \leq 1$, and we can therefore focus on approximating π over the set $\mathbb{I}$. Furthermore, let $\mathcal{A}(U)$ denote the greedy subset obtained by considering elements in $\mathbb{I}$ in decreasing order of density and picking the largest prefix where every element has nonnegative marginal profit. The following lemma implies that $\mathcal{A}(U)$ or the single best element is a 3-approximation to O^*.

Lemma 2. $\pi(O^*) \leq \pi(F^*) + \max_{e \in U} \pi(e) \leq \pi(\mathcal{A}(U)) + 2\max_{e \in U} \pi(e)$. *Therefore the integral greedy offline algorithm (Algorithm 2 in Appendix B) achieves a 3-approximation for $(\pi, 2^U)$.*

The offline greedy algorithm suggests an online solution as well. In the case where a single element gives a good approximation, we can use the classical secretary algorithm to get a good competitive ratio. In the other case, in order to get good competitive ratio, we merely need to estimate the smallest density, say ρ^-, in the prefix of elements that the offline greedy algorithm picks, and then accept every element that exceeds this threshold.

We pick an estimate for ρ^- by observing the first few elements of the stream U. Note that it is important for our estimate of ρ^- to be no smaller than ρ^-. In particular, if there are many elements with density just below ρ^-, and our algorithm uses a density threshold less than ρ^-, then the algorithm may be fooled into mostly picking elements with density below ρ^- (since elements arrive in random order), while the optimal solution picks elements with densities far exceeding ρ^-. We now describe how to pick an overestimate of ρ^- which is not too conservative, that is, there is still sufficient profit in elements whose densities exceed the estimate.

In the remainder of this section, we assume that every element has profit at most $\frac{1}{k_1+1}\pi(O^*)$ for an appropriate constant k_1, to be defined later. (If this does not hold, the classical secretary algorithm obtains an expected profit of at least $\frac{1}{e(k_1+1)}\pi(O^*)$). Then Lemma 2 implies $\pi(F^*) \geq \left(1 - \frac{1}{(k_1+1)}\right)\pi(O^*)$, $\max_{e\in U} \pi(e) \leq \frac{1}{k_1}\pi(F^*)$, and $\pi(\mathcal{A}(U)) \geq \left(1 - \frac{1}{k_1}\right)\pi(F^*)$.

We divide the stream U into two parts X and Y, where X is a random subset of U. Our algorithm unconditionally rejects elements in X and extracts a density threshold τ from this set. Over the remaining stream Y, it accepts an element if and only if its density is at least τ and if it brings in strictly positive marginal profit. Under the assumption of small element profits we can apply Lemma 1 to show that $\pi(X \cap \mathcal{A}(U))$ is concentrated and is a large enough fraction of $\pi(O^*)$. This implies that with high probability $\pi(X \cap \mathcal{A}(U))$ (which is a prefix of $\mathcal{A}(X)$) is a significant fraction of $\pi(\mathcal{A}(X))$. Therefore we attempt to identify $X \cap \mathcal{A}(U)$ by looking at profits of prefixes of X.

Algorithm 1 Online algorithm for single-dimensional $(\pi, 2^U)$

With probability 1/2 run the classic secretary algorithm to pick the single most profitable element else execute the following steps.
Draw k from Binomial$(n, 1/2)$.
Select the first k elements to be in the sample X. Unconditionally reject these elements.
Let τ be largest density such that $\pi(P_\tau^X) \geq \beta\left(1 - \frac{1}{k_1}\right)\pi(F^*(X))$ for constants β and k_1 to be specified later.
Initialize selected set $O \leftarrow \emptyset$.
for $i \in Y = U \setminus X$ **do**
  if $\pi(O \cup \{i\}) - \pi(O) \geq 0$ and $\rho(i) \geq \tau$ and $i \notin \mathbb{F}$ **then**
    $O \leftarrow O \cup \{i\}$
  else
    Exit loop.
  end if
end for

We will need the following lemma about $\mathcal{A}()$.

Lemma 3. *For any set S, consider subsets $A_1, A_2 \subseteq \mathcal{A}(S)$. If $A_1 \supseteq A_2$, then $\pi(A_1) \geq \pi(A_2)$. In other words, π is monotone-increasing when restricted to $\mathcal{A}(S)$ for all $S \subset U$.*

We define two good events. E_1 asserts that $X \cap \mathcal{A}(U)$ has high enough profit. Our final output is the set P_τ^Y. E_2 asserts that the profit of P_τ^Y is a large enough fraction of the profit of P_τ. Recall that $\mathcal{A}(U)$ is a density prefix, say P_{ρ^-}, and so $X \cap \mathcal{A}(U) = P_{\rho^-}^X$. We define the event E_1 as follows.

$$E_1: \ \pi(P_{\rho^-}^X) > \beta\ \pi(P_{\rho^-})$$

where β is a constant to be specified later. Conditioned on E_1, we have $\pi(P_{\rho^-}^X) > \beta\,(1 - 1/k_1)\,\pi(F^*) \geq \beta\,(1 - 1/k_1)\,\pi(F^*(X))$. Note that threshold τ, as selected by Algorithm 1, is the largest density such that $\pi(P_\tau^X) \geq \beta\,(1 - 1/k_1)\,\pi(F^*(X))$. Therefore, E_1 implies $\tau \geq \rho^-$, and we have the following lemma.

Lemma 4. *Conditioned on E_1, $O = P_\tau \cap Y \subseteq \mathcal{A}(U)$.*

On the other hand, $P_\tau^X \subseteq P_\tau \subset \mathcal{A}(U)$ along with Lemma 3 implies

$$\pi(P_\tau) \geq \pi(P_\tau^X) \geq \beta\,(1-1/k_1)\,\pi(F^*(X)) \geq \beta\,(1-1/k_1)\,\pi(P_{\rho^-}^X) \geq \beta^2\,(1-1/k_1)^2\,\pi(F^*)$$

where the second inequality is by the definition of τ, the third by optimality and the last is obtained by applying E_1 and $\mathcal{A}(U) \geq (1-1/k_1)\,F^*$.

We define ρ^+ to be the largest density such that $\pi(P_{\rho^+}) \geq \beta^2\left(1-\frac{1}{k_1}\right)^2\pi(F^*)$. Then $\rho^+ \geq \tau$, which implies $P_{\rho^+} \subseteq P_\tau$ and the following lemma.

Lemma 5. *Event E_1 implies $O \supseteq Y \cap P_{\rho^+}$.*

Based on the above lemma, we define event E_2 for an appropriate constant β' as follows

$$E_2 : \pi(P_{\rho^+}^Y) \geq \beta' \pi(P_{\rho^+})$$

Conditioned on events E_1 and E_2, and using Lemma 3 again, we get

$$\pi(O) \geq \pi(P_{\rho^+}^Y) \geq \beta'\beta^2(1-1/k_1)^2\pi(F^*)$$

To wrap up the analysis, we show that E_1 and E_2 are high probability events.

Lemma 6. *If no element of U has profit more than $\frac{1}{113}\pi(O^*)$, then $Pr[E_1 \wedge E_2] \geq 0.52$, where $\beta = 0.262$ and $\beta' = 0.094$.*

Putting everything together we get the following theorem.

Theorem 7. *Algorithm 1 achieves a competitive ratio of 616 for $(\pi, 2^U)$ using $k_1 = 112$ and $\beta = 0.262$.*

Proof. If there exists an element with profit at least $\frac{1}{113}\pi(O^*(U))$, the classical secretary algorithm (Step 1) gives a competitive ratio of $\frac{1}{113e} \geq \frac{1}{308}$. Otherwise, using Lemma 6, with $\beta' = 0.094$, we have $\mathbb{E}[\pi(O)] \geq \mathbb{E}[\pi(O) \mid E_1 \wedge E_2]\,\Pr[E_1 \wedge E_2] \geq 0.52\beta'\beta^2(1-1/k_1)^2\pi(F^*) \geq 0.52\beta'\beta^2(1-1/k_1)^2\,(1-1/(k_1+1))\,\pi(O^*) \geq \frac{1}{307}\pi(O^*)$. Since we flip a fair coin to decide whether to output the result of running the classical secretary algorithm, or output the set O, we achieve a $2\max\{308, 307\} = 616$-approximation to $\pi(O^*)$ in expectation (over the coin flip). □

4 Matroid-Constrained Profit Maximization

We now extend the algorithm of Section 3 to the setting $(\pi, \mathcal{F})$ where $\mathcal{F}$ is a matroid constraint. In particular, $\mathcal{F}$ is the set of all independent sets of a matroid over U. We skip a precise definition of matroids and will only use the following facts: $\mathcal{F}$ is a downward closed feasibility constraint and there exists an exact offline and an $O(\log r)$ online algorithm for $(\Phi, \mathcal{F})$, where Φ is a sum-of-values objective and r is the rank of the matroid. The algorithms and detailed proofs for this section are given in section C.

In the unconstrained setting, we showed that there always exists either a density prefix or a single element with near-optimal profit. So in the online setting it sufficed to determine the density threshold for a good prefix. In constrained settings this is no longer true, and we need to develop new techniques. Our approach is to develop a general reduction from the π objective to two different sum-of-values type objectives over the same feasibility constraint. This allows us to employ previous work on the $(\Phi, \mathcal{F})$ setting; we lose only a constant factor in the competitive ratio. We will first describe the reduction in the offline setting and then extend it to the online algorithm using techniques from Section 3.

Decomposition of π. For a given density γ, we define the *shifted density function* $h_\gamma()$ over sets as $h_\gamma(A) := \sum_{e\in A} (\rho(e) - \gamma)\, s(e)$ and the *fixed density function* $g_\gamma()$ over sizes as $g_\gamma(s) := \gamma s - \mathsf{C}(s)$. For a set A we use $g_\gamma(A)$ to denote $g_\gamma(s(A))$. It is immediate that for any density γ we can split the profit function as $\pi(A) = h_\gamma(A) + g_\gamma(A)$. In particular $\pi(O^*) = h_\gamma(O^*) + g_\gamma(O^*)$. Our goal will be to optimize the two parts separately and then return the better of them.

Note that the function h_γ is a sum of values function where the value of an element is defined to be $(\rho(e) - \gamma)s(e)$. Its maximizer is a subset of P_γ, the set of elements with nonnegative shifted density $\rho(e) - \gamma$. In order to ensure that the maximizer, say A, of h_γ also obtains good profit, we must ensure that $g_\gamma(A)$ is nonnegative, and therefore $\pi(A) \geq h_\gamma(A)$. This is guaranteed for a set A as long as $s(A) \leq \bar{\mathsf{s}}(\gamma)$.

Likewise, the function g_γ increases as a function of size s as long as s is at most $\bar{\mathsf{s}}(\gamma)$, and decreases thereafter. Therefore, in order to maximize g_γ, we merely need to find the largest (in terms of size) feasible subset of size no more than $\bar{\mathsf{s}}(\gamma)$. As before, if we can ensure that for such a subset h_γ is nonnegative (e.g. if the set is a subset of P_γ), then the profit of the set is no smaller than its g_γ value. This motivates the following definition of "bounded" subsets:

Definition 1. *Given a density γ a subset $A \subseteq U$ is said to be γ-bounded if $A \subseteq P_\gamma$ and $s(A) \leq \bar{\mathsf{s}}(\gamma)$.*

Proposition 8. *For any γ-bounded set A, $\pi(A) \geq h_\gamma(A)$ and $\pi(A) \geq g_\gamma(A)$.*

For a density γ and set T we define H_γ^T and G_γ^T as follows. (We write H_γ for H_γ^U and G_γ for G_γ^U.)

$$H_\gamma^T = \operatorname*{argmax}_{H\in\mathcal{F}, H\subset P_\gamma^T} h_\gamma(H) \qquad\qquad G_\gamma^T = \operatorname*{argmax}_{G\in\mathcal{F}, G\subset \bar{P}_\gamma^T} s(G)$$

Following our observations above, both H_γ and G_γ can be determined efficiently (in the offline setting) using standard matroid maximization. However, we must ensure that the two sets are γ-bounded. Further, in order to compare the performance of G_γ against O^*, we must ensure that its size is at least a constant fraction of the size of O^*. We now show that there exists a density γ for which H_γ and G_γ satisfy these properties.

Once again, we focus on the case where no single element has high enough profit by itself. Recall that F^* denotes $F^*(\mathbb{I})$, s^* denotes the size of this set and $\bar{P}_\gamma$ denotes $P_\gamma \setminus \{e_\gamma\}$.

Definition 2. *Let ρ^- be the largest density such that P_{ρ^-} has a feasible set of size at least s^*.*

We now state a useful property of ρ^-.

Lemma 9. *Any feasible set in $\bar{P}_{\rho^-}$ is ρ^--bounded and has size less than s^*. Moreover for any density $\gamma > \rho^-$ all feasible subsets of P_γ are γ-bounded.*

The following is our main claim of this section.

Lemma 10. *For any density $\gamma > \rho^-$, $\pi(O^*(\bar{P}_\gamma)) \leq \pi(H_\gamma) + \pi(G_\gamma)$. Furthermore, $\pi(O^*) \leq \pi(H_{\rho^-}) + \pi(G_{\rho^-}) + 2\max_{e\in U}\pi(e)$.*

This lemma immediately gives us an offline approximation for $(\pi, \mathcal{F})$: for every element density γ, we find the sets H_γ and G_γ; we then output the best (in terms of profit) of these sets or the best individual element. We obtain the following theorem:

Theorem 11. *Algorithm 3 (in Appendix C) 4-approximates $(\pi, \mathcal{F})$ in the offline setting.*

The online setting. Our online algorithm, as in the unconstrained case, uses a sample X from U to obtain an estimate τ for the density ρ^-. Then with equal probability it applies the online algorithm for $(h_\tau, \mathcal{F})$ on the remaining set $Y \cap P_\tau$ or the online algorithm for $(s, \mathcal{F})$ (in order to maximize g_τ) on $Y \cap P_\tau$. The algorithm is described in detail in Appendix C. Lemma 10 indicates that it should suffice for τ to be larger than ρ^- while ensuring that $\pi(O^*(P_\tau))$ is large enough. As in Section 3 we define the density ρ^+ as the upper limit on τ, and claim that τ satisfies the required properties with high probability.

Definition 3. *For a fixed parameter $\beta \leq 1$, let ρ^+ be the highest density with $\pi(O^*(P_{\rho^+})) \geq (\beta/16)^2 \pi(O^*)$.*

Lemma 12. *For fixed parameters $k_1 \geq 1$, $k_2 \leq 1$, $\beta \leq 1$ and $\beta' \leq 1$ suppose that there is no element with profit more than $\frac{1}{k_1}\pi(O^*)$. Then with probability at least k_2, we have that τ, as defined in Algorithm 4, satisfies $\rho^+ \geq \tau \geq \rho^-$ and $\pi(O^*(P_\tau^Y)) \geq \beta'(\beta/16)^2\pi(O^*)$.*

To conclude the analysis, we show that if the online algorithms for $(h_\tau, \mathcal{F})$ and $(s, \mathcal{F})$ have a competitive ratio of α, then we obtain an $O(\alpha)$ approximation to $\pi(O^*(P_\tau^Y))$. We therefore get the following theorem.

Theorem 13. *If there exists an α-competitive algorithm for the matroid secretary problem $(\Phi, \mathcal{F})$ where Φ is a sum-of-values objective, then Algorithm 4 in Appendix C achieves a competitive ratio of $O(\alpha)$ for the problem $(\pi, \mathcal{F})$.*

5 Multi-dimensional Profit Maximization

In this section, we consider the GSP with a multi-dimensional profit objective. Recall that in this setting each element e has ℓ different sizes $s_1(e), \ldots, s_\ell(e)$, and the cost of a subset is defined by ℓ different convex functions $\mathsf{C}_1, \ldots, \mathsf{C}_\ell$. The profit function is defined as $\pi(A) = v(A) - \sum_i \mathsf{C}_i(s_i(A))$.

As in the single-dimensional setting, we partition U into two sets $\mathbb{I}$ and $\mathbb{F}$ with $\mathbb{F} = \{e \in U : \operatorname{argmax}_\alpha \pi(\alpha e) < 1\}$. We first claim that as before an optimal solution cannot contain too many elements of $\mathbb{F}$. We therefore devote the remainder of this section to approximating π over $\mathbb{I}$. In this section, we focus on the unconstrained problem $(\pi, 2^U)$. Detailed algorithms and proofs are deferred to Appendix D. We defer the discussion of the constrained settings to Appendix E.

Lemma 14. *$|O^* \cap \mathbb{F}| \leq \ell$.*

The main challenge of this setting is that we cannot summarize the value-size tradeoff that an element provides by a single density because the element can be quite large in one dimension and very small in another. Our high level approach is to distribute the value of each element across the ℓ dimensions, thereby defining densities and decomposing profit across dimensions appropriately. We do this in such a way that a maximizer of the ith dimensional profit for some dimension i gives us a good overall solution (albeit at a cost of a factor of ℓ).

Formally, let $\rho : U \to \mathbb{R}^\ell$ denote an ℓ-dimensional vector function $\rho(e) = (\rho_1(e), \ldots, \rho_\ell(e))$ that satisfies $\sum_i \rho_i(e)s_i(e) = v(e)$ for all e. We set $v_i(e) = \rho_i(e)s_i(e)$ and $\pi_i(A) = v_i(A) - \mathsf{C}_i(s_i(A))$ and note that $\pi(A) = \sum_i \pi_i(A)$. Let F_i^* denote the maximizer of π_i over $\mathbb{I}$. Then, $\pi(F^*) \leq \sum_i \pi_i(F_i^*)$.

Given this observation, it is natural to try to obtain an approximation to π by solving for F_i^* for all i and rounding the best one. This does not immediately work: even if $\pi_i(F_i^*)$ is very large,

$\pi(F_i^*)$ could be negative because of the profit of the set being negative in other dimensions. We will now describe an approach for defining and finding density vectors such that the best set F_i^* indeed gives an $O(\ell)$ approximation to $O^*(\mathbb{I})$. We first define a quantity $\Pi_i(\gamma_i)$ which bounds the ith dimensional profit that can be obtained by any set with elements of ith density at most γ_i: $\Pi_i(\gamma_i) = \max_t(\gamma_i t - \mathsf{C}_i(t))$. We can bound $\pi_i(F_i^*)$ using this quantity.

Lemma 15. *For a given density γ_j, let $A = \{a \in F_j^* : \rho_j(a) \geq \gamma_j\}$. Then $\pi_j(F_j^*) \leq \pi_j(A) + \Pi_j(\gamma_j)$.*

In order to obtain a uniform bound on the profits $\pi_j(F_j^*)$, we restrict density vectors as follows. We call a vector $\rho(e)$ *proper* if it satisfies the following properties:

(P1) $\sum_i \rho_i(e) s_i(e) = v(e)$

(P2) $\Pi_i(\rho_i(e)) = \Pi_j(\rho_j(e))$ for all $i, j \in [1, \ell]$; we denote this quantity by $\Pi(e)$.

The following lemma is proved in Appendix D.2.

Lemma 16. *For every element e, a unique proper density vector exists and can be found in polynomial time.*

Finally, we note that proper density vectors induce a single ordering over elements. In particular, since the Π_is are monotone, $\rho_i(e) \geq \rho_i(e')$ if and only if $\Pi(e) \geq \Pi(e')$. We order the elements $e_1, \ldots, e_n$ in decreasing order of Π. Note that each F_i^* is a (fractional) prefix of this sequence. Let F_1^* be the shortest prefix. Let $\mathcal{A} = \{e_1, \ldots, e_{k_1}\}$ denote the integral part of F_1^* and $e' = e_{k_1+1}$(*i.e.*, F_1^*'s unique fractional element if it exists). We get the following lemma by noting that $\pi_1(F_1^*) \geq \Pi_1(e')$(see Lemma 25 in Appendix D).

Lemma 17. *For proper $\rho(e)$s and $\mathcal{A}$ and e' as defined above, for every i, $\pi_i(F_i^*) \leq \pi_i(\mathcal{A}) + \pi_1(F_1^*) \leq 2\pi(F_1^*)$. Furthermore, $\pi(F^*) \leq \ell(\pi(\mathcal{A}) + \pi(e'))$.*

Lemmas 14 and 17 together give $\pi(O^*) \leq \ell(\pi(\mathcal{A}) + 2\max_e \pi(e))$, and therefore imply an offline 3ℓ-approximation for $(\pi, 2^U)$ in the multi-dimensional setting.

The online setting. Note that proper densities essentially define a 1-dimensional manifold in ℓ-dimensional space. We can therefore hope to apply our online algorithm from Section 3 to this setting. However, there is a caveat: the algorithm from Section 3 uses the offline algorithm as a subroutine on the sample X to estimate the threshold τ; naïvely replacing the subroutine by the $O(\ell)$ approximation described above leads to an $O(\ell^2)$ competitive online algorithm[3]. In order to improve the competitive ratio to $O(\ell)$ we need to pick the threshold τ more carefully.

We define τ to be the largest density with $e_\tau \in X$ such that for an appropriate constant β, $\pi(\bar{P}_\tau^X) + \pi(e_\tau) \geq \frac{\beta}{2}\pi_i(F_i^*(X))$ for all i.

For a set T, let $F_i^*(T)$ denote the maximizer of π_i over $T \cap \mathbb{I}$ and let $\mathrm{P}^*(T) = \cap_i F_i^*(T)$ denote the shortest of these prefixes. Recall that $\mathrm{P}^*(U) = F_1^*$. Let ρ^- denote the smallest density in F_1^*. That is, $F_1^* = \mathrm{P}^*(P_{\rho^-})$. Our analysis relies on the following two events:

$$E_1 : \pi(\mathrm{P}^*(P_{\rho^-}^X)) \geq \beta\pi(\mathrm{P}^*(P_{\rho^-}))$$
$$E_1' : \pi(O^*(X)) \geq \beta''\pi(O^*).$$

[3]Note the $(1 - 1/k_1)^2$ factor in the final competitive ratio in Theorem 7; this factor is due to the use of the offline subroutine in determining τ.

E_1 implies the following sequence of inequalities; here the second inequality follows from Lemma 17.

$$\pi(\mathrm{P}^*(P^X_{\rho^-})) \geq \beta\pi(F^*_1) \geq \frac{\beta}{2}\pi_i(F^*_i) \geq \frac{\beta}{2}\pi_i(F^*_i(X)). \tag{1}$$

This implies $\tau \geq \rho^-$. (See Lemma 26 in Appendix D for a formal statement and proof.) Summing over all dimensions and applying event E'_1 gives us

$$\ell(\pi(\bar{P}^X_\tau) + \pi(e_\tau)) \geq \frac{\beta}{2}\sum_i \pi_i(F^*_i(X)) \geq \frac{\beta}{2}\sum_i \pi_i(O^*(X)) = \frac{\beta}{2}\pi(O^*(X)) \geq \frac{\beta\beta''}{2}\pi(O^*)$$

So if we define ρ^+ to be the highest density such that $\pi(\bar{P}_{\rho^+}) + \pi(e_{\rho^+}) \geq \frac{\beta\beta''}{2\ell}\pi(O^*)$, then we get $\rho^+ \geq \tau$. Then, as before we can define the event E_2 in terms of ρ^+ to conclude that $\pi(P^Y_\tau)$ is large enough. The last part of our analysis shows that the events E_1, E'_1, and E_2 hold with constant probability.

Lemma 18. *Suppose that* $\max_e \pi(e) \leq (1/k_3\ell)\pi(O^*)$. *Then,* $\Pr[E_1 \wedge E'_1 \wedge E_2] \geq 0.28$.

Via a similar argument as for Theorem 7, we get

Theorem 19. *Algorithm 5 in Appendix D is $O(\ell)$ competitive for $(\pi, 2^U)$ where π is a multi-dimensional profit function.*

6 Discussion

A natural question suggested by our work is whether we can obtain a good competitive ratio for $(\pi, \mathcal{F})$ or even $(\pi, 2^U)$ when π is an arbitrary non-montone submodular function (that is, not necessarily a nonnegative one). The hardness for multi-dimensional knapsack indicates that this is not possible. Nevertheless it would be interesting to explore the GSP for more general objectives that can take on negative values. A more immediate open problem is to obtain an $O(\alpha\ell)$ competitive algorithm for the ℓ-dimensional $(\pi, \mathcal{F})$ problem where α is the competitive ratio of an online algorithm for $\mathcal{F}$ with a sum-of-values objective.

Acknowledgements

We thank Jason Hartline for suggesting the use of the balanced sampling lemma from [11].

References

[1] M. Babaioff, M. Dinitz, A. Gupta, N. Immorlica, and K. Talwar. Secretary problems: weights and discounts. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1245–1254. Society for Industrial and Applied Mathematics, 2009.

[2] M. Babaioff, J. Hartline, and R. Kleinberg. Selling banner ads: Online algorithms with buyback. In *Fourth Workshop on Ad Auctions*, 2008.

[3] M. Babaioff, J. D. Hartline, and R. D. Kleinberg. Selling ad campaigns: Online algorithms with cancellations. In *Proceedings of the tenth ACM conference on Electronic commerce*, pages 61–70. ACM, 2009.

[4] M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg. A knapsack secretary problem with applications. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 16–28. Springer Berlin / Heidelberg, 2007.

[5] M. Babaioff, N. Immorlica, and R. Kleinberg. Matroids, secretary problems, and online mechanisms. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 434–443. Society for Industrial and Applied Mathematics, 2007.

[6] M. H. Bateni, M. T. Hajiaghayi, and M. Zadimoghaddam. Submodular secretary problem and extensions. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 39–52. Springer Berlin / Heidelberg, 2010.

[7] S. Chawla, J. D. Hartline, D. L. Malec, and B. Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the 42nd ACM symposium on Theory of computing*, STOC '10, pages 311–320. ACM, 2010.

[8] B. C. Dean, M. X. Goemans, and J. Vondrák. Adaptivity and approximation for stochastic packing problems. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '05, pages 395–404. Society for Industrial and Applied Mathematics, 2005.

[9] N. B. Dimitrov and C. G. Plaxton. Competitive weighted matching in transversal matroids. In *Proceedings of the 35th international colloquium on Automata, Languages and Programming, Part I*, pages 397–308. Springer-Verlag, 2008.

[10] E. B. Dynkin. The optimum choice of the instant for stopping a Markov process. *Soviet Math. Dokl*, 4(627-629), 1963.

[11] U. Feige, A. Flaxman, J. D. Hartline, and R. Kleinberg. On the Competitive Ratio of the Random Sampling Auction. In *Proceedings of the first annual workshop on Internet and network economics*. Springer, 2005.

[12] T. Ferguson. Who solved the secretary problem. *Statist. Sci.*, 4(3):282–289, 1989.

[13] P. R. Freeman. The secretary problem and its extensions: a review. *International Statistical Review*, 51(2):189–206, 1983.

[14] Anupam Gupta, Aaron Roth, Grant Schoenebeck, and Kunal Talwar. Constrained non-monotone submodular maximization: Offline and secretary algorithms. In *Internet and Network Economics*, pages 246–257. Springer Berlin / Heidelberg, 2010.

[15] D. P. Kennedy. Prophet-type inequalities for multi-choice optimal stopping. *Stochastic Processes and their Applications*, 24(1):77–88, 1987.

[16] R. Kleinberg. A multiple-choice secretary algorithm with applications to online auctions. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 630–631. Society for Industrial and Applied Mathematics, 2005.

[17] N. Korula and M. Pál. Algorithms for secretary problems on graphs and hypergraphs. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming: Part II*, ICALP '09, pages 508–520. Springer-Verlag, 2009.

[18] J. Oxley. *Matroid Theory.* Oxford University Press, 1992.

[19] E. Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *The Annals of Probability*, 12(4):1213–1216, 1984.

[20] S. Samuels. Secretary problems. In *Handbook of Sequential Analysis*, pages 381–405. Marcel Dekker, 1991.

[21] J. A. Soto. Matroid Secretary Problem in the Random Assignment Model. In *Proceedings of the twenty-second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1275–1284. Society for Industrial and Applied Mathematics, 2011.

A Balanced Sampling

In this section, we give a proof of Lemma 1. We begin with a restatement of Lemma 1 from [11] since it plays a crucial role in our argument. Note that we choose a different parameterization than they do, since in our setting the balance between approximation ratio and probability of success is different.

Lemma 20. *Let X_i, for $i \geq 1$ be indicator random variables for a sequence of independent, fair coinflips. Then, for $S_i = \sum_{k=1}^{i} X_k$, we have* $\Pr[\forall i,\ S_i \geq \lfloor i/3 \rfloor] \geq 0.76$.

We now proceed to prove Lemma 1. While we do not give a closed form for the approximation factor $\beta(c)$ in the statement of the lemma, we define it implicitly as

$$\beta(c) = \max_{0<y<1} \left(\frac{y}{2+y}\right)\left(1 - \frac{2}{c(1-y)}\right),$$

and give explicit values in Table 1 for particular values of c that we invoke the lemma with.

Proof of Lemma 1. Our general approach will be to separate our set of weights I into a "good" set G and a "bad" set B. At a high level, Lemma 20 guarantees us that at worst, we will accept a weight every third time we flip a coin. So the good case is when weights do not decrease too quickly; this intuition guides our definitions of G and B.

Let $y \in (0,1)$ be a lower bound on "acceptable" rates of decrease; we tune the exact value later. Throughout, we use $w(S)$, $\underline{w}(S)$, and $\overline{w}(S)$ to denote the total, minimum, and maximum weights of a set S. We form G and B as follows.

Initialize $B = \emptyset$ and $i = 1$. Consider I in order of decreasing weight. While $I \neq \emptyset$, we repeat the following. Let P be the largest prefix such that $\underline{w}(P) \geq y \cdot \overline{w}(P)$. If $|P| \leq 2$, move P from I to B, i.e. set $B := B \cup P$ and $I := I \setminus P$. Otherwise, $|P| \geq 3$ and we define G_i to be the 3 largest elements of P; remove them from I (i.e. set $I := I \setminus G_i$); and increment i by 1. Once we are done, we define $G = \cup_i G_i$.

First, we show that the total weight in B cannot be too large. Note that we add at most 2 elements at a time to B; and when we do add elements, we know that all remaining elements in I (and hence all later additions to B) are smaller by more than a factor of y. Thus, we can see that

$$w(B) \leq \sum_{i \geq 0} 2y^i \cdot \overline{w}(B) \leq \frac{2\overline{w}(B)}{1-y} \leq \frac{2w(I)}{c(1-y)},$$

c given	y chosen	$\beta(c)$ achieved
111	0.84	≈ 0.262
15/2	0.46	≈ 0.094
8	.47	≈ 0.100

Figure 1: Some specific values of approximation ratio achieved by Lemma 1

.

by our assumption that no individual weight is more than $w(I)/c$.

Next, we show that with probability at least 0.76, we can lower bound the fraction of weight we keep from G. Consider applying Lemma 20, flipping coins first for the weights in G in order by decreasing weight. Note that by the time we finish flipping coins for G_i, we must have added at least i weights to J; hence the i^{th} weight we add to J must have value at least $\underline{w}(G_i)$. On the other hand, we know that

$$w(G_i) \leq 2\overline{w}(G_i) + \underline{w}(G_i) \leq \left(\frac{2}{y} + 1\right) \underline{w}(G_i),$$

and so summing over i we can see that elements we accept have total weight $w(J) \geq (\frac{y}{2+y})w(G)$.

Combining our bounds for the weights of G and B, we can see that with probability 0.76 the elements we accept have weight

$$w(J) \geq \left(\frac{y}{2+y}\right) w(G) = \left(\frac{y}{2+y}\right)(w(I) - w(B)) \geq \left(\frac{y}{2+y}\right)\left(1 - \frac{2}{c(1-y)}\right) w(I);$$

optimizing the above with respect to y for a fixed c gives the claimed result. Note that for each fixed $y \in (0, 1)$ our approximation factor is increasing in c, and so the optimal value $\beta(c)$ must be increasing in c as well. □

B Algorithms and proofs for the unconstrained setting (Section 3)

In this section we present a greedy approximation algorithm for the unconstrained offline settings, and give the omitted proofs from Section 3. Throughout, we use $\mathcal{A}(S)$ to denote the subset of $\mathbb{I}$ which is selected by Algorithm 2 when it is applied to S.

Proof of Lemma 2. We first show that O^* has at most one element from $\mathbb{F}$. For contradiction, assume that O^* contains at least two elements $f_1, f_2 \in \mathbb{F}$. Since densities are unique, without loss of generality, we assume $\rho(f_1) > \rho(f_2)$.

Recall that $\mathbb{F}$ is precisely the set of elements for which it is optimal to accept a strictly fractional amount. Let $\alpha_1, \alpha_2 < 1$ be the optimal fractions for f_1 and f_2, i.e. $\text{argmax}_\alpha \pi(\alpha f_1) = \alpha_1$ and $\text{argmax}_\alpha \pi(\alpha f_2) = \alpha_2$. Then adding a fractional amount of any element with density at most $\rho(f_1)$ to $\{\alpha_1 f_1\}$ results in strictly decreased profit. But this implies $\pi(O^*) < \pi(O^* \setminus \{f_2\})$, contradicting the optimality of O^*.

Let f be the unique element in $O^* \cap \mathbb{F}$. By subadditivity, we get that $\pi(O^* \setminus \{f\}) + \pi(f) \geq \pi(O^*)$. Since $O^* \setminus \{f\} \subseteq \mathbb{I}$, we have $\pi(F^*) \geq \pi(O^* \setminus \{f\})$. In the rest of the proof we focus on approximating $\pi(F^*)$.

Algorithm 2 Offline algorithm for single-dimensional $(\pi, 2^U)$

Initialize set $\mathbb{I} \leftarrow \{a \in U \mid \rho(a) \geq \mathsf{c}(s(a))\}$
Initialize selected set $\mathcal{A} \leftarrow \emptyset$
Sort $\mathbb{I}$ in decreasing order of density.
for $i \in \mathbb{I}$ **do**
  if marginal profit for the ith element, $\pi_{\mathcal{A}}(i) \geq 0$ **then**
    $\mathcal{A} \leftarrow \mathcal{A} \cup \{i\}$
  else
    Exit loop.
  end if
end for
Set $m := \text{argmax}_{\{a \in U\}} \pi(a)$ {the most profitable element}
if $\pi(m) > \pi(\mathcal{A})$ **then**
  Output set $\{m\}$
else
  Output set $\mathcal{A}$
end if

Note that F^* is a fractional density prefix of $\mathbb{I}$. So let $F^* = P_{\rho^-} \cup \{\alpha e\}$, for some e and $\alpha < 1$. The subadditivity of π implies $\pi(F^*) \leq \pi(P_{\rho^-}) + \pi(\{\alpha e\})$. Note that Algorithm 2 selects $\mathcal{A} = P_{\rho^-}$, and that $\pi(\{\alpha e\}) \leq \pi(e)$ since $e \in \mathbb{I}$.

Hence, combining the above inequalities we get $\pi(\mathcal{A}) + \pi(e) + \pi(f) \geq \pi(O^*)$. This in turn proves the required claim. □

Proof of Lemma 3. We observe that the fractional greedy algorithm sorts its input S by decreasing order of density, and $\mathcal{A}(S)$ consists of the top $|\mathcal{A}(S)|$ elements under that ordering. Since $F^*(S)$ contains each element in $\mathcal{A}(S)$ in its entirety, we can see that $F^*(B) = B$ for any subset B of $\mathcal{A}(S)$. So for $A_2 \subseteq A_1 \subseteq \mathcal{A}(S)$, we have that $F^*(A_2) = A_2 \subseteq A_1 = F^*(A_1)$; by the optimality of F^*, this implies that $\pi(A_2) \leq \pi(A_1)$ as claimed. □

Proof of Lemma 6. We show that $\Pr[E_1], \Pr[E_2] \geq 0.76$; the desired inequality $\Pr[E_1 \wedge E_2] \geq 0.52$ then follows by the union bound.

In the following, we assume the elements are sorted in decreasing order of density. Denote the profit increment of the i^{th} element by $\tilde{\pi}(i) = \pi(P_{(i)}) - \pi(P_{(i-1)})$; this extends naturally to sets $A \subseteq \mathcal{A}(U)$ by setting $\tilde{\pi}(A) = \sum_{i \in A} \tilde{\pi}(i)$. By the subadditivity of π, we have $\pi(A) \geq \tilde{\pi}(A)$ for all $A \subseteq \mathcal{A}(U)$, with equality at $\mathcal{A}(U)$.

We apply Lemma 1 with P_{ρ^-} as the fixed set I and $P^X_{\rho^-} = U \cap P_{\rho^-}$ as the the random set J. The weights in Lemma 1 correspond to profit increments of the elements. Note that $P_{\rho^-} = \mathcal{A}(U)$; so we know both that elements in P_{ρ^-} have non-negative profit increments, and $\pi(P_{\rho^-}) \geq \pi(O^*) - 2\max_{e \in U} \pi(e)$. Hence if no element has profit exceeding a $1/113$-fraction of $\pi(O^*)$, we get that any element $e_i \in P_{\rho^-}$ has profit increment $\tilde{\pi}(i) \leq \pi(e_i) \leq 1/(113-2)\pi(P_{\rho^-}) = 1/111\tilde{\pi}(P_{\rho^-})$. Hence we can apply Lemma 1 to get $\tilde{\pi}(P^X_{\rho^-}) \geq 0.262\tilde{\pi}(P_{\rho^-})$ with probability at least 0.76, and so the event E_1 holds with probability at least 0.76.

By our definition of ρ^+, the profit of P_{ρ^+} is at least $\beta^2(1-1/k_1)^2(1-1/(k_1+1))\pi(O^*)$; substituting in the specified values of β and k_1 give us that no element in P_{ρ^+} has profit increment more than $2/15\ \tilde{\pi}(P_{\rho^+})$. Thus, applying Lemma 1 implies $\Pr[E_2] \geq 0.76$ with $\beta' = 0.094$. □

C Algorithms and proofs for the constrained setting (Section 4)

C.1 The offline setting

We begin by formally proving that the function g_γ increases as a function of size s as long as s is at most $\bar{\mathsf{s}}(\gamma)$.

Lemma 21. *If density γ and sizes s and t satisfy $s \leq t \leq \bar{\mathsf{s}}(\gamma)$, then $g_\gamma(s) \leq g_\gamma(t)$.*

Proof. Since $\mathsf{C}(\cdot)$ is convex, we have that its marginal, $\mathsf{c}()$, is monotonically non-decreasing. Thus we get the following chain of inequalities,

$$\mathsf{C}(t) - \mathsf{C}(s) = \sum_{z=s+1}^{t} \mathsf{c}(z) \leq (t-s) \times \mathsf{c}(t) \leq (t-s)\gamma.$$

The last inequality follows from the assumption that t is no more than $\bar{\mathsf{s}}(\gamma)$ and hence $\mathsf{c}(t) \leq \gamma$. By definition of $g_\gamma()$ we get the desired claim. □

Proof of Proposition 8. Since $\pi(A) = h_\gamma(A) + g_\gamma(A)$, it is sufficient to prove that $h_\gamma(A)$ and $g_\gamma(A)$ are both non-negative. The former is clearly non-negative since all elements of A have density at least γ. Lemma 21 implies that the latter is non-negative by taking $t = s(A)$ and $s = 0$. □

Before we proceed we need the following fact about the fractional optimal subset F^*. Recall that s^* denotes the size of F^*.

Lemma 22. *If F^* has an element of density γ then s^* is at most $\bar{\mathsf{s}}(\gamma)$.*

Proof. The proof is by contradiction. Say s^* is more than $\bar{\mathsf{s}}(\gamma)$. Recall that e_γ denotes the element with density γ. We show that in such conditions reducing the fractional contribution of e_γ, say by ϵ, increases profit giving us a better fractional solution. This is turn contradicts the optimality of F^*.

Write $s = s(e_\gamma)$ and note that

$$\begin{aligned}\pi(F^*) &= v(F^*) - \mathsf{C}(s^*)\\ &= [(v(F^*) - \gamma \times \epsilon s) - \mathsf{C}(s^* - \epsilon s)] - [\mathsf{C}(s^*) - \mathsf{C}(s^* - \epsilon s) - \gamma \times \epsilon s]\,.\end{aligned}$$

If ϵ is such that $s^* - \epsilon > \bar{\mathsf{s}}(\gamma)$, then we have $\mathsf{C}(s^*) - \mathsf{C}(s^* - \epsilon s) > \gamma \times \epsilon s$; thus the term $[\mathsf{C}(s^*) - \mathsf{C}(s^* - \epsilon s) - \gamma \times \epsilon s]$ is positive which proves the claim.

□

Proof of Lemma 9. By definition, the size of any feasible set contained in $P_{\rho^-} \setminus \{e_{\rho^-}\}$ is no more than s^*.

We will show that $s^* \leq \bar{\mathsf{s}}(\rho^-)$. Then the first part of the lemma follows immediately. For the second part we have $\gamma > \rho^-$ and hence $\bar{\mathsf{s}}(\rho^-) \leq \bar{\mathsf{s}}(\gamma)$. Overall a size bound of s^* also implies that feasible sets in P_γ satisfy the size requirement for being γ-bounded. Hence we get that all feasible sets in $P_{\rho^-} \setminus \{e_{\rho^-}\}$ are ρ^--bounded and all feasible sets in P_γ are γ-bounded.

The size of F^* is at most the size of its support. Thus the support of F^* is a feasible set of size at least s^*. By definition, ρ^- is the largest density such that P_{ρ^-} contains a feasible set of size s^*. Hence we get that F^* contains an element of density less than or equal to ρ^-. Applying Lemma 22 we get $s^* \leq \bar{\mathsf{s}}(\rho^-)$ and the lemma follows. □

Proof of Lemma 10. Let P, H, and G denote P_{ρ^-}, H_{ρ^-}, and G_{ρ^-} respectively. As in the unconstrained setting, there can be at most one element in the intersection of O^* and $\mathbb{F}$ (see proof of Lemma 2). Note that $\pi()$ is subadditive, hence $\pi(O^* \cap \mathbb{I}) + \max_{e \in U} \pi(e) \geq \pi(O^*)$. In the analysis below we do not consider elements present in $\mathbb{F}$ and show that $\pi(H) + \pi(G) + \max_{e \in U} \pi(e) \geq \pi(O^* \cap \mathbb{I})$. This in turn establishes the second part of the Lemma.

For ease of notation we denote e_{ρ^-} as e'. Without loss of generality, we can assume that H does not contain e' since $\rho(e') - \rho^- = 0$. Therefore set H is contained in $P \setminus \{e'\}$. By Lemma 9 we get that H is ρ^--bounded.

Note that G does not contain e', by definition, hence its size is at most s^*. Also, $G \cup \{e'\}$ is no smaller than maximum-size feasible subset contained in P. So, by definition of ρ^-, we also have $s(G) + s(e') \geq s^*$. Thus there exists $\alpha < 1$ such that the fractional set $F = G \cup \{\alpha e'\}$ has size exactly equal to s^*.

Next we split the profit of F^* into two parts and bound the first by $h_{\rho^-}(H)$ and the second by $g_{\rho^-}(F)$.

$$\pi(F^*) = h_{\rho^-}(F^*) + g_{\rho^-}(s^*)$$

Note that we can drop elements which have a negative contribution to the sum to get

$$\begin{aligned} h_{\rho^-}(F^*) &\leq h_{\rho^-}(F^* \cap P_{\rho^-}) \\ &\leq h_{\rho^-}(\mathrm{supp}(F^*) \cap P_{\rho^-}) \\ &\leq h_{\rho^-}(H) \\ &\leq \pi(H) \end{aligned}$$

We can only increase the value of a subset by "rounding up" fractional elements so we get the second inequality. The third inequality follows from the optimality of H, and the fourth from the fact that it is ρ^--bounded.

To bound the second part we note that $s(F) = s^*$, hence $g_{\rho^-}(F) = \rho^- s^* - \mathsf{C}(s^*)$. Elements in F have density no less than ρ^- and its size is bounded above by $\bar{\mathsf{s}}(\rho^-)$, hence it is a ρ^--bounded set implying that $\pi(F) \geq g_{\rho^-}(F)$. Note that $F = G \cup \{\alpha e'\}$, by sub-additivity of $\pi()$ we have $\pi(G) + \pi(\alpha e') \geq \pi(F)$. Moreover $e' \in \mathbb{I}$ implies $\pi(\alpha e') \leq \pi(e')$ and hence we get

$$\pi(F^*) \leq \pi(H) + g_\gamma(F) \leq \pi(H) + \pi(G) + \pi(e')$$

which proves the second half of the lemma.

The first half of the lemma follows along similar lines. We have the standard decomposition, $\pi(O^*(\bar{P}_\gamma)) = h_\gamma(O^*(\bar{P}_\gamma)) + g_\gamma(O^*(\bar{P}_\gamma))$. By definition, H_γ is the constrained maximizer of h_γ, hence we get $h_\gamma(O^*(\bar{P}_\gamma)) \leq h_\gamma(H_\gamma)$. We note that all feasible sets in P_γ are γ-bounded, for density $\gamma > \rho^-$ (Lemma 9). Hence, by Lemma 21, g_γ strictly increases with size when restricted to feasible sets in $\bar{P}_\gamma$. G_γ is the largest such set, hence we get $g_\gamma(G_\gamma) \geq g_\gamma(O^*(\bar{P}_\gamma))$. H_γ and G_γ are γ-bounded and hence by Proposition 8 we have $\pi(H_\gamma) \geq h_\gamma(H_\gamma)$ and $\pi(G_\gamma) \geq g_\gamma(G_\gamma)$. This establishes the lemma. □

Algorithm 3 Offline algorithm for single-dimensional $(\pi, \mathcal{F})$

1: Set $A_{\max} \leftarrow \operatorname{argmax}_{H \in \{H_\gamma\}_\gamma} \pi(H)$
2: Set $B_{\max} \leftarrow \operatorname{argmax}_{G \in \{G_\gamma\}_\gamma} \pi(G)$
3: Set $e_{\max} \leftarrow \operatorname{argmax}_{e \in U} \pi(e)$
4: Assign $\mathcal{A}(U) \leftarrow \operatorname{argmax}_{S \in \{A_{\max}, B_{\max}, e_{\max}\}} \pi(S)$

C.2 The online setting

The online algorithm (Algorithm 4) is described in detail below. Note that we use an algorithm for $(\Phi, \mathcal{F})$ where Φ is a sum-of-values objective in Algorithm 4 as a black box. For example if the underlying feasibility constraint is a partition matroid we execute the partition matroid secretary algorithm in steps 11 and 20. Since the algorithm we use is an online algorithm, we can execute steps 11 and 20 in parallel with the respective 'for' loops in steps 13 and 22. This ensures that all accepted elements have positive profit increment.

Proof of Lemma 12. We first define two events analogous to the events E_1 and E_2 in Section 3.

$$E_1 : \pi(O^*(P^X_{\rho^-})) \geq \beta\pi(O^*(P_{\rho^-}))$$
$$E_2 : \pi(O^*(P^Y_{\rho^+})) \geq \beta'\pi(O^*(P_{\rho^+}))$$

The next two claims use event E_1 to imply $\rho^+ \geq \tau \geq \rho^-$. Hence we get the following containment $P^Y_{\rho^+} \subset P^Y_\tau \subset P^Y_{\rho^-}$, which implies $\pi(O^*(P^Y_\tau)) \geq \pi(O^*(P^Y_{\rho^+}))$. This inequality along with event E_2 and the definition of ρ^+ proves the second condition. Finally we show that the probability of E_1 and E_2 occurring simultaneously is at least k_2.

Claim 1. *Event E_1 implies that $\rho^+ \geq \tau$ and so $P^Y_{\rho^+} \subset P^Y_\tau$.*

Proof. First, by containment and optimality, we observe that

$$\pi(O^*(P_\tau)) \geq \pi(O^*(P^X_\tau)) \geq \pi(\mathcal{A}(P^X_\tau)).$$

By definition of τ, we have $\pi(\mathcal{A}(P^X_\tau)) \geq \frac{\beta}{16}\pi(\mathcal{A}(X))$. Furthermore,

$$\pi(\mathcal{A}(X)) \geq \frac{1}{4}\pi(O^*(X)) \geq \frac{1}{4}\pi(O^*(P^X_{\rho^-}))$$

Lemma 10 gives us $\pi(O^*(P_{\rho^-})) \geq \frac{1}{4}\pi(O^*)$. This together with event E_1 implies

$$\pi(O^*(P^X_{\rho^-})) \geq \beta\pi(O^*(P_{\rho^-})) \geq \frac{\beta}{4}\pi(O^*).$$

Thus, we have that $\pi(O^*(P_\tau)) \geq \left(\frac{\beta}{16}\right)^2 \pi(O^*)$. We have defined ρ^+ to be the largest density for which the previous profit inequality holds. Hence we conclude that $\rho^+ \geq \tau$. □

Claim 2. *Event E_1 implies that $\tau \geq \rho^-$.*

Algorithm 4 Online algorithm for single-dimensional $(\pi, \mathcal{F})$

Draw k from Binomial$(n, 1/2)$.
Select the first k elements to be in the sample X. Unconditionally reject these elements.
Toss a fair coin.
if Heads **then**
  Set output O as the first element, over the remaining stream, with profit higher than $\max_{e \in X} \pi(e)$.
else
  Determine $\mathcal{A}(X)$ using the offline Algorithm 3.
  Let β be a specified constant and let τ be the highest density such that $\pi(\mathcal{A}(P_\tau^X)) \geq \frac{\beta}{16}\pi(\mathcal{A}(X))$.
  Toss a fair coin.
  if Heads **then**
    Let O_1 be the result of executing an online algorithm for $(h_\tau, \mathcal{F})$ on the subset P_τ^Y of the remaining stream with the objective function

$$h_\tau(A) = \sum_{e \in A} (\rho(e) - \tau) s(e)$$

    Set $O \leftarrow \emptyset$.
    for $e \in O_1$ **do**
      if $\pi(O \cup \{e\}) - \pi(O) \geq 0$ **then**
        Set $O \leftarrow O \cup \{e\}$.
      end if
    end for
    Output O.
  else
    Let O_2 be the result of executing an online algorithm for $\mathcal{F}$ on the subset P_τ^Y of the remaining stream with objective function $s()$.
    Set $O \leftarrow \emptyset$.
    for $e \in O_2$ **do**
      if $\pi(O \cup \{e\}) - \pi(O) \geq 0$ **then**
        Set $O \leftarrow O \cup \{e\}$.
      end if
    end for
    Output O.
  end if
end if

Proof. As stated before, Lemma 10 gives us $\pi(O^*(P_{\rho^-})) \geq \frac{1}{4}\pi(O^*)$. We have that $\pi(\mathcal{A}(P^X_{\rho^-})) \geq \frac{1}{4}\pi(O^*(P^X_{\rho^-}))$. Now, E_1 implies that

$$\pi(O^*(P^X_{\rho^-})) \geq \beta\pi(O^*(P_{\rho^-})) \geq \frac{\beta}{4}\pi(O^*).$$

Combining these we get that

$$\pi(\mathcal{A}(P^X_{\rho^-})) \geq \frac{\beta}{16}\pi(O^*) \geq \frac{\beta}{16}\pi(\mathcal{A}(X)).$$

Since τ is the largest density for which the above inequality holds we have $\tau \geq \rho^-$. □

Finally, we state a lower bound on the probability that E_1 and E_2 occur simultaneously. We observe that $\pi(O^*(P^X_{\rho^-})) \geq \pi(O^*(P_{\rho^-}) \cap X)$, and so it suffices to bound the probability that $\pi(O^*(P_{\rho^-}) \cap X) \geq \beta\pi(O^*(P_{\rho^-}))$ and likewise that $\pi(O^*(P_{\rho^+}) \cap Y) \geq \beta'\pi(O^*(P_{\rho^+}))$.

Claim 3. *For a fixed constant k_1, if no element of S has profit more than $\frac{1}{k_1}\pi(O^*)$ then $Pr[E_1 \wedge E_2] \geq 0.52$*

Proof. The proof of this claim is similar to that of Lemma 6. We show that $\Pr[E_1] \geq 0.76$ and $\Pr[E_2] \geq 0.76$; the result then follows by applying the union bound.

We fix an ordering over elements of $O^*(P_{\rho^-})$, such that the profit increments are non-increasing. That is, if L_i is the set containing elements 1 through i and hence the profit increment of the ith element is, $\tilde{\pi}(i) := \pi(L_i) - \pi(L_{i-1})$ then we have $\tilde{\pi}(1) \geq \tilde{\pi}(2) \geq \ldots$. Note that such an ordering can be determined by greedily picking elements from $O^*(P_{\rho^-})$ such that the profit increment at each step is maximized.

We set $k_1 \geq 24$, now the fact that no element in S has profit more than $\frac{1}{24}\pi(O^*)$ implies no element by itself has profit more than $1/8\ \pi(O^*(P_{\rho^-}))$, since $\pi(O^*(P_{\rho^-})) \geq 1/3\pi(O^*)$. Profit increments of elements are upper bounded by their profits, therefore we can apply Lemma 1 with $O^*(P_{\rho^-})$ as the fixed set and profit increments as the weights. By optimality of $O^*(P_{\rho^-})$ we have that these profit increments are non-negative hence the required conditions of Lemma 1 hold and we have $\pi(O^*(P_{\rho^-}) \cap X) \geq \beta\pi(O^*(P_{\rho^-}))$ with probability at least 0.76. Since $\pi(O^*(P^X_{\rho^-})) \geq \pi(O^*(P_{\rho^-}) \cap X)$ we get $\Pr[E_1] \geq 0.76$ with $\beta = 1/10$.

By definition of ρ^+ the profit of $O^*(P_{\rho^+})$ is at least $\left(\frac{\beta}{9}\right)^2 \pi(O^*)$. With $k_1 \geq 8\left(\frac{9}{\beta}\right)^2$ and $\beta' = 1/10$ we can again Lemma 1 to show that $\Pr[E_2] \geq 0.76$. Hence we can set $k_2 = 0.52$ and this completes the proof of the claim. □

The three claims together imply the lemma. □

Before we proceed to prove Theorem 13, we show that in steps 9 to 27 the algorithm obtains a good approximation to $O^*(P^Y_\tau)$.

Lemma 23. *Suppose that there is an α-competitive algorithm for $(\Phi, \mathcal{F})$ where Φ is any sum-of-values objective. For a fixed set Y and threshold τ, satisfying $\tau \geq \rho^-$, we have $\mathrm{E}_\sigma[\pi(O_1) + \pi(O_2)] \geq \alpha\ \pi(O^*(P^Y_\tau))$, where the expectation is over all permutations σ of Y.*

Proof. The threshold τ is either equal to or strictly greater than ρ^{-}. In the former case $e_{\rho^{-}}$ must have been in the sample set X and hence $O_1, O_2 \subseteq \bar{P}_{\rho^{-}}$. By Lemma 9 we show that O_1 and O_2 are ρ^{-}-bounded and hence τ-bounded. On the other hand if $\tau > \rho^{-}$ we can again apply Lemma 9 and get that O_1 and O_2 are τ-bounded.

Hence by Proposition 8 we get following inequalities $\mathrm{E}_\sigma[\pi(O_1)] \geq \mathrm{E}_\sigma\left[h_\tau(O_1)\right]$ and $\mathrm{E}_\sigma[\pi(O_2)] \geq \mathrm{E}_\sigma[g_\tau(O_2)]$.

By applying the α-competitive matroid secretary algorithm with objective h_τ (Step 11 of Algorithm 4) we get

$$\begin{aligned}
\mathrm{E}_\sigma[h_\tau(O_1)] &\geq \alpha \times h_\tau(H_\tau^Y) \\
&\geq \alpha \times h_\tau(O^*(P_\tau^Y))
\end{aligned}$$

The second inequality follows from the optimality of H_τ^Y.

Next we bound $\mathrm{E}_\sigma[g_\tau(O_2)]$. Let K be the largest feasible subset contained in P_τ^Y. The fact that the underlying algorithm is α-competitive implies $\mathrm{E}_\sigma[s(O_2)] \geq \alpha \times s(K)$.

Note that, as observed above, $P_\tau^Y \subset \bar{P}_{\rho^{-}}$. Since $K \subset P_\tau^Y$, by definition of ρ^{-} we get that $s(K) \leq s^*$. So, for O_2 we have

$$\begin{aligned}
\mathrm{E}_\sigma[g_\tau(O_2)] &\geq \mathrm{E}_\sigma[\tau s(O_2) - \mathsf{C}(s(O_2)] \\
&\geq \mathrm{E}_\sigma\left[\tau\left(\frac{s(O_2)}{s(K)}\right)s(K) - \left(\frac{s(O_2)}{s(K)}\right)\mathsf{C}(s(K))\right] \\
&= \mathrm{E}_\sigma\left[\frac{s(O_2)}{s(K)}\right](\tau s(K) - \mathsf{C}(s(K))) \\
&\geq \alpha(\tau s(K) - \mathsf{C}(s(K))) \\
&= \alpha\ g_\tau(K) \\
&\geq \alpha\ g_\tau(O^*(P_\tau^Y))
\end{aligned}$$

Since $s(O^*(P_\tau^Y)) \leq s(K) \leq s^* \leq \bar{\mathsf{s}}(\tau)$ we get the last inequality by applying Lemma 21.

The conclusion of the lemma now follows from the decomposition, $\pi(O^*(P_\tau^Y)) = h_\tau(O^*(P_\tau^Y)) + g_\tau(O^*(P_\tau^Y))$. $\square$

Proof of Theorem 13. With probability $\frac{1}{2}$ we apply the standard secretary algorithm which is e-competitive. If an element has profit more than $\frac{1}{k_1}\pi(O^*)$, in expectation we get a profit of $\frac{1}{2k_1e}$ times the optimal.

We have $\Pr[E_1 \wedge E_2] \geq k_2$, for a fixed constant k_2. Also, the events E_1 and E_2 depend only on what elements are in X and Y, and not on their ordering in the stream. So conditioned on E_1 and E_2, the remaining stream is still a uniformly random permutation of Y. Therefore, if no element has profit more than $\frac{1}{k_1}\pi(O^*)$ we can apply the second inequality of Lemma 12 and Lemma 23 along with the fact that we output O_1 and O_2 with probability $\frac{1}{4}$ each, to show that

$$\begin{aligned}
\mathrm{E}[\pi(O)] &\geq \frac{1}{4}\mathrm{E}[\pi(O_1) + \pi(O_2) \mid E_1 \wedge E_2] \times \Pr[E_1 \wedge E_2] \\
&\geq \frac{\alpha k_2}{4}\mathrm{E}[\pi(O^*(P_\tau^Y)) \mid E_1 \wedge E_2] \\
&\geq \frac{\alpha k_2 \beta'}{4}\left(\frac{\beta}{16}\right)^2 \pi(O^*).
\end{aligned}$$

Overall, we have that

$$\mathrm{E}[\pi(O)] \geq \min\left\{\frac{\alpha k_2 \beta'}{4}\left(\frac{\beta}{16}\right)^2, \frac{1}{2k_1 e}\right\} \cdot \pi(O^*)$$

Since all the involved parameters are fixed constants we get the desired result. ☐

D Algorithms and proofs for the multi-dimensional unconstrained setting (Section 5)

In this section we present the omitted proofs from Section 5. First, we deal with $O^* \cap \mathbb{F}$ by proving Lemma 14.

Proof of Lemma 14. Suppose, towards a contradiction, that $|O^* \cap \mathbb{F}| \geq \ell + 1$. For $i \in \{1, \dots, \ell\}$, let m_i be any element in O^* with $s_i(m_i) = \max_{e \in O^*} s_i(e)$. Since $|O^* \cap \mathbb{F}| \geq \ell + 1$, there exists $o \in (O^* \cap \mathbb{F}) \setminus \{m_1, \dots, m_\ell\}$ such that $s_i(o) \leq s_i(O^* \setminus \{o\})$ for all i. This implies that when we compare the marginal cost of adding another copy of o to $\{o\}$ against the marginal cost of adding o to $O^* \setminus \{o\}$, we have by convexity

$$\sum_i \mathsf{C}_i(s_i(o+o)) - \mathsf{C}_i(s_i(o)) \leq \sum_i \mathsf{C}_i(s_i(O^*)) - \mathsf{C}_i(s_i(O^* \setminus \{o\}))$$

Therefore, we have $\pi(O^*) - \pi(O^* \setminus \{o\}) \leq \pi(o+o) - \pi(o) < 0$ since $o \in \mathbb{F}$, and this contradicts the optimality of O^*. ☐

Next, we restrict our attention to $\mathbb{I}$ and prove Lemma 17. Recall that $\Pi_i(\gamma_i) = \max_t(\gamma_i t - \mathsf{C}_i(t))$ and note that $\bar{\mathsf{s}}_i(\gamma_i) = \mathrm{argmax}_t(\gamma_i t - \mathsf{C}_i(t))$ is the size t at which $\Pi_i(\gamma_i)$ is achieved. First, we need the following fact about F_1^*. It implies that the multidimensional profit function π is monotone when restricted to subsets of F_1^*.

Lemma 24. *Consider subsets $A_1, A_2 \subset F_1^*$. If $A_1 \supset A_2$, then $\pi(A_1) \geq \pi(A_2)$. Furthermore, $\pi(A) \geq \pi_i(A)$ for all i if $A \subset F_1^*$.*

Proof. For any i, we have $F_1^* \subset F_i^*$. Since F_i^* is the fractional prefix that optimizes $\pi_i(\cdot)$, applying Lemma 3 to $\pi_i(\cdot)$ implies that $\pi_i(A_1) \geq \pi_i(A_2) \geq 0$. By summing, we have $\pi(A_1) \geq \pi(A_2)$, and also that $\pi(A_1) \geq \pi_i(A_1)$. Setting A_1 as A gives us the second claim. ☐

We now prove Lemma 15.

Proof of Lemma 15. Subadditivity implies $\pi_j(F_j^*) \leq \pi_j(A) + \pi_j(F_j^* \setminus A)$. Since the elements e in $F_j^* \setminus A$ have $\rho_j(e) \leq \gamma_j$, we have $\pi_j(F_j^* \setminus A) \leq \max_t \gamma_j t - \mathsf{C}_j(t) = \Pi_j(\gamma_j)$. ☐

We observe that Lemma 15 gives us a bound on $\pi_j(F_j^*)$ that is in terms of a prefix A and Π_j. The following lemma allows us to exploit (P2) to obtain a bound on Π_j in terms of $\pi_1(F_1^*)$.

Lemma 25. *If F_i^* does not contain e or contains αe for $\alpha < 1$, then $\Pi_i(e) \leq \pi_i(F_i^*)$.*

Proof. Let $\rho'_i = \min_{a \in F_i^*} \rho_i(a)$. We observe that $\rho_i(e) \leq \mathsf{c}_i(s_i(F_i^*))$ since otherwise, we could have included an additional fractional amount of e to F_i^* and increased its profit. Together with Lemma 22, we have $\bar{\mathsf{s}}_i(e) \leq s_i(F_i^*) \leq \bar{\mathsf{s}}_i(\rho'_i)$. Thus, applying Lemma 21 and the fact that $\rho'_i \geq \rho_i(e)$, we have

$$\Pi_i(e) = \rho_i(e)\bar{\mathsf{s}}_i(e) - \mathsf{C}_i(\bar{\mathsf{s}}_i(e)) \leq \rho'_i\bar{\mathsf{s}}_i(e) - \mathsf{C}_i(\bar{\mathsf{s}}_i(e)) \leq \rho'_i s_i(F_i^*) - \mathsf{C}_i(s_i(F_i^*)).$$

Since F_i^* is ρ'_i-bounded in the ith dimension, we have $\pi_i(F_i^*) \geq \rho'_i s_i(F_i^*) - \mathsf{C}_i(s_i(F_i^*))$ by Lemma 8 and this completes the proof. □

We now have the necessary ingredients to prove Lemma 17.

Proof of Lemma 17. Since ρ is proper, we have $\Pi_i(e') = \Pi_1(e')$. Recall that e' is not in $\mathcal{A}$, the integral subset of F_1^*, so Lemma 25 implies $\Pi_i(e') = \Pi_1(e') \leq \pi_1(F_1^*)$. Together with Lemma 15, this gives us

$$\pi_i(F_i^*) \leq \pi_i(\mathcal{A}) + \pi_1(F_1^*) \tag{2}$$

Now, Lemma 24 gives us $\pi(F_1^*) \geq \pi_i(F_1^*) \geq \pi_i(\mathcal{A})$ and $\pi(F_1^*) \geq \pi_1(F_1^*)$, so we have $\pi_i(F_1^*) \leq 2\pi(F_1^*)$ as claimed.

Summing Equation (2) over i and applying subadditivity to $\pi_1(F_1^*)$, we have

$$\begin{aligned}\sum_i \pi_i(F_i^*) &\leq \sum_{i \neq 1} [\pi_i(\mathcal{A}) + \pi_1(F_1^*)] + \pi_1(F_1^*) \\ &\leq \sum_{i \neq 1} \pi_i(\mathcal{A}) + \ell(\pi_1(\mathcal{A}) + \pi_1(\alpha e')) \\ &= \pi(\mathcal{A}) + (\ell - 1)\pi_1(\mathcal{A}) + \ell\pi_1(\alpha e').\end{aligned}$$

Applying Lemma 24 we have $\pi(\mathcal{A}) \geq \pi_1(\mathcal{A})$ and $\pi(\alpha e') \geq \pi_1(\alpha e')$, and so

$$\pi(\mathcal{A}) + (\ell - 1)\pi_1(\mathcal{A}) + \ell\pi_1(\alpha e') \leq \ell(\pi(\mathcal{A}) + \pi(\alpha e')).$$

The conclusion now follows from the fact we are considering only elements from $\mathbb{I}$ and $\pi(F^*) = \sum_i \pi_i(F^*) \leq \sum_i \pi_i(F_i^*)$. □

D.1 Multi-dimensional Unconstrained Online

In this subsection, we give the online algorithm for unconstrained multidimensional profit maximization and the rest of the omitted proofs from Section 5.

Lemma 26. *Conditioned on E_1, we have $\tau \geq \rho^-$.*

Proof. Since $\bar{P}^X_{\rho^-} \subset \bar{P}_{\rho^-} = \mathcal{A}$, Lemma 24 implies that $\bar{P}^X_{\rho^-} \subset \mathrm{P}^*(P^X_{\rho^-})$. Event E_1 guarantees that $P^X_{\rho^-}$ is non-empty, so let ρ' be the minimum density of $P^X_{\rho^-}$. We have $P^X_{\rho'} = P^X_{\rho^-}$, which implies $\bar{P}^X_{\rho'} \subset \bar{P}^X_{\rho^-} \subset \mathrm{P}^*(P^X_{\rho^-})$ and $P^X_{\rho^-} = P^X_{\rho'} = \bar{P}^X_{\rho'} \cup \{e_{\rho'}\}$ by definition. So, we can write $\mathrm{P}^*(P^X_{\rho^-}) = \mathrm{P}^*(P^X_{\rho'}) = \bar{P}^X_{\rho'} \cup \{\alpha' e_{\rho'}\}$ for some $\alpha' \in [0, 1]$. Using subadditivity and the fact that $e_{\rho'} \in \mathbb{I}$ gives us

$$\pi(\bar{P}^X_{\rho'}) + \pi(e_{\rho'}) \geq \pi(\bar{P}^X_{\rho'} \cup \{\alpha' e_{\rho'}\}) = \pi(\mathrm{P}^*(P^X_{\rho^-})). \tag{3}$$

Algorithm 5 Online algorithm for multi-dimensional $(\pi, 2^U)$

1: With probability 1/2 run the classic secretary algorithm to pick the single most profitable element else execute the following steps.
2: Draw k from Binomial$(n, 1/2)$.
3: Select the first k elements to be in the sample X. Unconditionally reject these elements.
4: Let τ be largest density such that $e_\tau \in X$ satisfies $\pi(\bar{P}_\tau^X) + \pi(e_\tau) \geq \frac{\beta}{2}\pi_i(F_i^*(X))$ for all i, for a constant β.
5: Initialize selected set $O \leftarrow \emptyset$.
6: **for** $i \in Y = U \setminus X$ **do**
7:   **if** $\pi(O \cup \{i\}) - \pi(O) \geq 0$ and $\rho(i) \geq \tau$ and $i \notin \mathbb{F}$ **then**
8:     $O \leftarrow O \cup \{i\}$
9:   **else**
10:     Exit loop.
11:   **end if**
12: **end for**

Recall that event E_1 implies Equation (1), which we restate here:

$$\pi(\mathrm{P}^*(P_{\rho^-}^X)) \geq \frac{\beta}{2}\pi_i(F_i^*(X)) \quad (\forall i). \tag{1}$$

Now, threshold τ as selected by Algorithm 5 is the largest density such that $\pi(\bar{P}_\tau^X) + \pi(e_\tau) \geq \frac{\beta}{2}\pi(F_i^*(X))$ for all i. Since step 4 of the algorithm would have considered $\pi(\bar{P}_{\rho'}^X) + \pi(e_{\rho'})$, Equations (3) and (1) imply $\tau \geq \rho'$. By definition, $\rho' \geq \rho^-$, so we conclude that E_1 implies $\tau \geq \rho^-$. □

Furthermore, by definition, $e_\tau \in X$ and so $e_\tau \notin P_\tau^Y$, which implies that $P_\tau^Y \subset \bar{P}_\tau$. Since $\tau \geq \rho^-$ implies $\bar{P}_\tau \subset \bar{P}_{\rho^-} = \mathcal{A}$, Lemma 26 gives us the following lemma.

Lemma 27. *Conditioned on E_1, $O = P_\tau^Y \subset \bar{P}_\tau \subset \mathcal{A}$.*

Next, we formally prove the following statement.

Lemma 28. *Conditioned on E_1 and E_1', $O = P_\tau^Y \supset P_{\rho^+}^Y$.*

Proof. Recall that from Equation (1) and event E_1', we have $\ell(\pi(\bar{P}_\tau^X) + \pi(e_\tau)) \geq \frac{\beta\beta''}{2}\pi(O^*)$. We have $\bar{P}_\tau^X \subset \bar{P}_\tau \subset \mathcal{A}$ by Lemma 27, so Lemma 24 gives us $\pi(\bar{P}_\tau) \geq \pi(\bar{P}_\tau^X)$. Therefore, we have $\pi(\bar{P}_\tau) + \pi(e_\tau) \geq \frac{\beta\beta''}{2\ell}\pi(O^*)$. Since ρ^+ is the highest density such that $\pi(\bar{P}_{\rho^+}) + \pi(e_{\rho^+}) \geq \frac{\beta\beta''}{2\ell}\pi(O^*)$, we conclude that $\rho^+ \geq \tau$. □

Based on the above lemma, we define event E_2 for some fixed constant β' as follows.

$$E_2 : \pi(P_{\rho^+}^Y) \geq \beta'\pi(P_{\rho^+}).$$

Since $P_{\rho^+}^Y \subset P_\tau^Y \subset \bar{P}_\tau \subset \mathcal{A}$, applying Lemma 24, we have $\pi(O) = \pi(P_\tau^Y) \geq \pi(P_{\rho^+}^Y)$. Therefore, conditioning on events E_1, E_1' and E_2, we get

$$\pi(O) \geq \pi(P_{\rho^+}^Y) \geq \frac{\beta\beta''\beta'}{2\ell}\pi(O^*).$$

To wrap up the analysis, we argue that the probability of these events is bounded from below by a constant. Using Lemma 1 as in the proof of Lemma 6 , if no element has profit at least $\frac{1}{\ell k_3}\pi(O^*)$, then events E_1, E_1' and E_2 each occur with probability at least 0.76. Using a union bound, we have that $\Pr[E_1 \wedge E_1' \wedge E_2] \geq 0.28$. This proves Lemma 18 and implies Theorem 19.

D.2 Computing Proper Densities

In this subsection, we show that proper ρ's exist and can be efficiently computed. To do this, we make the following assumptions about the marginal cost functions:

(A1) The marginal cost functions $\mathsf{c}_j(\cdot)$ are unbounded.

(A2) They satisfy $\mathsf{c}_j(s_j(U)) > \max_{e \in U} v(e)$ for all i.

Note that these assumptions do not affect $C_i(t)$ for sizes $t \leq s_i(U)$, and so have no effect on either the profit function or the optimal solution as well. We first prove that there exists a proper $\rho(e)$ for each element e. Observe that (P1) and (P2) uniquely defines $\rho(e)$. Therefore, we only need to find x^* satisfying the equation

$$\sum_j \Pi_j^{-1}(x^*) s_j(e) = v(e); \tag{4}$$

then our proper density $\rho(e)$ is given by $\rho_j(e) = \Pi_j^{-1}(x^*)$. By Assumption (A1), Π_j is a strictly-increasing continuous function with $\Pi_j(0) = 0$ and $\lim_{\gamma \to \infty} \Pi_j(\gamma) = \infty$, so its inverse Π_j^{-1} is well-defined and is also a strictly-increasing continuous function with $\Pi_j^{-1}(0) = 0$ and $\lim_{x \to \infty} \Pi_j^{-1}(x) = \infty$. Thus, the solution x^* exists.

Next, we show that for any element e, we can efficiently compute $\rho(e)$. In the following, we fix an element e, and use s_j and v to denote $s_j(e)$ and $v(e)$, respectively. We define $I_j(t) = \Pi_j(\mathsf{c}_j(t))$. Let x^* be the solution to Equation 4 for the fixed element e and note that $x^* = \Pi(e)$.

In the following two lemma statements and proofs, we focus on a single dimension j and remove subscripts for ease of notation. First, we prove that we can easily compute $I(t)$.

Lemma 29. *We have $I(t) = \mathsf{c}(t)t - \mathsf{C}(t)$.*

Proof. Let $t' = \bar{\mathsf{s}}(\mathsf{c}(t))$. Since t' is the maximum size $r \geq t$ such that $\mathsf{c}(t) \geq \mathsf{c}(r)$, by monotonicity, we have that $\mathsf{c}(\cdot)$ is constant in $[t, t']$. This implies $\mathsf{C}(t') - \mathsf{C}(t) = \mathsf{c}(t)(t' - t)$ and so,

$$\mathsf{c}(t)t' - \mathsf{C}(t') = \mathsf{c}(t)t - \mathsf{C}(t).$$

The lemma now follows from the fact that $I(t) = \Pi(\mathsf{c}(t))$ is the LHS of the above equation. ☐

Next, we prove a lemma that helps us determine $\Pi^{-1}(x)$ given x.

Lemma 30. *Given x and positive integer t such that $I(t) \leq x < I(t+1)$, we have*

$$\Pi^{-1}(x) = \frac{x + \mathsf{C}(t)}{t}.$$

Proof. By definition of $I(\cdot)$, we have $\Pi(\mathsf{c}(t)) \le x < \Pi(\mathsf{c}(t+1))$, and since $\Pi(\cdot)$ is strictly increasing, we get

$$\mathsf{c}(t) \le \Pi^{-1}(x) < \mathsf{c}(t+1).$$

By definition of $\bar{\mathsf{s}}(\cdot)$, this gives us $t \le \bar{\mathsf{s}}(\Pi^{-1}(x)) < t+1$, and therefore $\bar{\mathsf{s}}(\Pi^{-1}(x)) = t$, since $\mathsf{c}(\cdot)$ changes only on the integer points.

Thus, we have

$$x = \Pi(\Pi^{-1}(x)) = \Pi^{-1}(x)t - \mathsf{C}(t)$$

and solving this for $\Pi^{-1}(x)$ gives us the desired equality. □

Lemma 30 leads to the FIND-DENSITY algorithm which, given a profit x, uses a binary search to compute $\Pi_j^{-1}(x)$. Together with Lemma 31, this enables us to determine x^* by first using binary search, and then solving linear equations.

Algorithm 6 Given x and sizes s_j, find $\Pi^{-1}(x)$ and $t_1, \ldots, t_\ell$ satisfying $I_j(t_j) \le x < I_j(t_j+1)$.

FIND-DENSITY(x, s)

```
1: for j = 1 to ℓ do
2:    Binary search to find integral t_j ∈ [0, s_j(U)) satisfying I_j(t_j) ≤ x < I_j(t_j + 1).
3:    Set ρ_j ← (x + C_j(t_j))/t_j.
4: end for
5: return (ρ, t), where ρ is the vector (ρ_1, ..., ρ_ℓ) and t is the vector (t_1, ..., t_ℓ).
```

Lemma 31. *Suppose we have a positive integer x such that*

$$\sum_j \Pi_j^{-1}(x)s_j \le v < \sum_j \Pi_j^{-1}(x+1)s_j, \tag{5}$$

and positive integers $t_1, \ldots, t_\ell$ such that $I_j(t_j) \le x < I_j(t_j+1)$ for all j. Then the solution x^* to Equation (4) is precisely:

$$x^* = \frac{v - \sum_j \mathsf{C}_j(t_j)(s_j/t_j)}{\sum_j (s_j/t_j)}.$$

Proof. Equation (5) and the monotonicity of the Π_j^{-1}'s imply that $x \le x^* < x+1$ and so, we have $I_j(t_j) \le x^* < x+1 \le I_j(t_j+1)$ for all j. Applying Lemma 30 to dimension j gives us $\Pi_j^{-1}(x^*) = (x^* - \mathsf{C}_j(t_j))/t_j$. Then, applying Equation (4), we get

$$v = \sum_j \Pi_j^{-1}(x^*)s_j = \sum_j \left[\frac{x^* + \mathsf{C}_j(t_j)}{t_j}\right] s_j.$$

Solving this for x^* gives the claimed equality. □

We need the following lemma to show that the binary search upper bounds in both algorithms are correct.

Lemma 32. *For proper density ρ, we have $x^* < I_j(s_j(U))$ for all dimensions j.*

Algorithm 7 Given sizes s_j and value v, find ρ satisfying (P1) and (P2).

FIND-PROPER-DENSITY(v, s)

Binary search to find integral $x \in [0, \min_j I_j(s_j(U)))$ satisfying

$$\sum_j \delta_j^- s_j \le v < \sum_j \delta_j^+ s_j,$$

where $(\delta^-, t) = $ FIND-DENSITY(x, s) and $(\delta^+, t') = $ FIND-DENSITY$(x+1, s)$.

Set

$$x^* \leftarrow \frac{v - \sum_j \mathsf{C}_j(t_j)(s_j/t_j)}{\sum_j (s_j/t_j)}.$$

for $j = 1$ to ℓ **do**

  Set $\rho_j \leftarrow (x^* + \mathsf{C}_j(t_j))/t_j$.

end for

return ρ

Proof. If an element has zero size in dimension j, then we can ignore $\mathsf{C}_j(\cdot)$. So without loss of generality, we assume that $s_j > 0$ for all j. Property (P1) gives us $v = \sum_i \rho_i s_i \ge \rho_j s_j$ and so $\rho_j \le v/s_j \le v$. From assumption (A2), we have that $\mathsf{c}_j(s_j(U)) > v \ge \rho_j$. This implies that

$$\Pi_j(\rho_j) < \Pi(\mathsf{c}_j(s_j(U))) = I_j(s_j(U))$$

Since ρ is proper, we have $\Pi_j(\rho_j) = x^*$ and this proves the lemma. □

Theorem 33. *Algorithms* FIND-DENSITY *and* FIND-PROPER-DENSITY *run in polynomial time and are correct.*

Proof. Lemmas 30 and 31 show that given correctness of the binary search upper bounds, the output is correct. Lemma 32 implies that the binary search upper bound of FIND-PROPER-DENSITY is correct. We observe that FIND-DENSITY is only invoked for integral profits $x < I_j(s_j(U))$. Therefore, we have $I_j(t') \le x < I_j(t'+1)$ for $t' < s_j(U)$ and this proves that the binary search upper bound of FIND-DENSITY is correct.

Since the numbers involved in the arithmetic and the binary search are polynomial in terms of $s_j(U)$, $\mathsf{C}_j(s_j(U))$, we conclude that the algorithms take time polynomial in the input size. □

The online setting. When the algorithm does not get to see the entire input at once, then it does not know $s_j(U)$. However, we can get around this by observing that if we have sizes $t_1, \ldots, t_\ell$ satisfying $I_j(t_j) > \Pi_j(\rho_j) = x^*$ for all j, then $I_j(t_j) > x^*$ and hence $\min_j I_j(t_j)$ suffices as an upper bound for the binary search in FIND-PROPER-DENSITY. Since we invoke FIND-DENSITY for $x < I_j(t_j)$, we have that $t_1, \ldots, t_\ell$ suffice as upper bounds for the binary searches in FIND-DENSITY as well.

By (A2), we have $\mathsf{c}_j(s_j(U)) > v \ge \rho_j$ for any proper ρ, so if we guess t_j such that $\mathsf{c}_j(t_j) > v$, then we have $I_j(t_j) > \Pi_j(\rho_j)$. Therefore, we set $t_j = 2^m$, with $m = 1$ initially, increment m one at a time and check if $\mathsf{c}_j(t_j) < v$. Assumption (A2) and monotonicity of $\mathsf{c}_j(\cdot)$ implies that $t_j \le 2s_j(U)$ and so it takes us $m \le \log(2s_j(U))$ iterations to get $\mathsf{c}_j(t_j) > v$. Repeating this for each dimension gives us sufficient upper bounds for the binary searches in both algorithms.

E Multi-dimensional costs with general feasibility constraint

In this section we consider the multi-dimensional costs setting with a general feasibility constraint, $(\pi, \mathcal{F})$. As before we use an $O(1)$-approximate offline and an α-competitive online algorithm for $(\Phi, \mathcal{F})$ as a subroutine, where Φ is a sum-of-values objective. While we will be able to obtain an $O(\alpha\ell)$ approximation in the offline setting, this only translates into an $O(\alpha\ell^5)$ competitive online algorithm.

As in Section 5, we associate with each element e an ℓ-dimensional proper density vector $\rho(e)$. Then we decompose the profit functions π_i into sum-of-values objectives defined as follows.

$$g^i_\gamma(A) = \gamma_i s_i(A) - \mathsf{C}_i(s_i(A)),$$

$$h^i_\gamma(A) = \sum_{e\in A}(\rho_i(e) - \gamma_i)s_i(e),$$

Let $h_\gamma(A) = \sum_i h^i_\gamma(A)$ and $g_\gamma(A) = \sum_i g^i_\gamma(A)$. We have $\pi_i(A) = h^i_\gamma(A) + g^i_\gamma(A)$ for all i and $\pi(A) = h_\gamma(A) + g_\gamma(A)$. As before, $G_{\gamma,i} = \operatorname{argmax}_{A\subset \bar{P}_\gamma} s_i(A)$.

We extend the definition of "boundedness" to the multi-dimensional setting as follows (see also Definition 1).

Definition 4. *Given a density vector γ, a subset $A \subset P_\gamma$ is said to be γ-bounded if, for all i, A is γ_i-bounded with respect to C_i, that is, $\rho_i(A) \geq \gamma_i$ and $s_i(A) \leq \bar{\mathsf{s}}_i(\gamma_i)$. If A is γ-bounded and γ is the minimum density of A, then we say A is* bounded.

The following lemma is analogous to Proposition 8.

Lemma 34. *If A is bounded, then $\pi(A) \geq \pi_i(A)$ for all i. Moreover, if A is γ-bounded then for all i, we have*

$$\pi(A) \geq g_\gamma(A) \geq g^i_\gamma(A),$$

$$\pi(A) \geq h_\gamma(A) \geq h^i_\gamma(A).$$

Proof. We start by assuming that A is γ-bounded. For each dimension i, we get $\pi_i(A) \geq g^i_\gamma(A) \geq 0$ by Proposition 8. Thus, we have $\pi(A) \geq g_\gamma(A) \geq g^i_\gamma(A)$ by summing over i. A similar proof shows that $\pi(A) \geq h_\gamma(A) \geq h^i_\gamma(A)$. Next, we assume that A is bounded. Let μ be its minimum density. Then $h^i_\mu(A) \geq 0$ and $g^i_\mu(A) \geq 0$. This gives us $\pi_i(A) = h^i_\mu(A) + g^i_\mu(A) \geq 0$ and so $\pi(A) \geq \pi_i(A)$ for all i. □

As in the single-dimensional setting, our approach is to find an appropriate density γ to bound $h_\gamma(O^*)$ and $g_\gamma(O^*)$ in terms of the maximizers of h^i_γ and g^i_γ. We use Algorithm 3, the offline algorithm for the single-dimensional constrained setting given in Section 4, as a subroutine. We consider two possible scenarios: either all feasible sets are bounded or there exists an unbounded set. We use the following lemma to tackle the first scenario.

Lemma 35. *Suppose all feasible sets are bounded. Let $D_i = \operatorname{argmax}_{D\in\{H_{\gamma_i}, G_{\gamma_i}, e\}_{\gamma_i}} \pi_i(D)$ be the result of running Algorithm 3 on the single-dimensional instance $(\pi_i, \mathcal{F})$, where $\mathcal{F}$ is the underlying feasibility constraint. Then, we have $\sum_i \pi(D_i) \geq \frac{1}{4}\pi(O^*)$.*

Proof. Fix a dimension i. Since all feasible sets are bounded, we have $\pi(D_i) \geq \pi_i(D_i)$ by Lemma 34. Furthermore, Theorem 3 implies that $\pi_i(D_i) \geq \frac{1}{4}\pi_i(O^*)$. Summing over all dimensions, we have

$$\sum_i \pi(D_i) \geq \sum_i \pi_i(D_i) \geq \sum_i \frac{1}{4}\pi_i(O^*) = \frac{1}{4}\pi(O^*). \qquad \square$$

Now we handle the case when there exists an unbounded feasible set.

Lemma 36. *Suppose there exists an unbounded feasible set. Let ρ^- be the highest proper density such that P_{ρ^-} contains an unbounded feasible set. Let D_i' be the result of running Algorithm 3 on $\bar{P}_{\rho^-}$ (i.e. ignoring elements of density at most ρ^-) with objective π_i and feasibility constraint $\mathcal{F}$. If $e_{\rho^-} \in \mathbb{I}$, then we have*

$$4\sum_i \pi(D_i') + \ell(\max_i \pi(G_{\rho^-,i}) + \pi(e_{\rho^-})) \geq \pi(O^*).$$

Proof. For brevity we use e' to denote e_{ρ^-}. Let $O_1^* = O^* \cap (\bar{P}_{\rho^-})$ and $O_2^* = O^* \setminus O_1^*$. By subadditivity, we have $\pi(O^*) \leq \pi(O_1^*) + \pi(O_2^*)$. We observe that all feasible sets in $\bar{P}_{\rho^-}$ are bounded by definition of ρ^-. Thus, we can apply Lemma 35 and get $4\sum_i \pi(D_i') \geq \pi(O_1^*)$.

Next, we approximate $\pi(O_2^*)$. We have $h_{\rho^-}(O_2^*) \leq 0$ since O_2^* has elements of density at most ρ^-. So, we have $\pi(O_2^*) \leq \sum_i g^i_{\rho^-}(O_2^*)$. We know that for all i,

$$\begin{aligned}\Pi(e') &= \max_t(\rho_i^- t - \mathsf{C}_i(t)) \\ &\geq \rho_i^- s_i(O_2^*) - \mathsf{C}_i(s_i(O_2^*)) \\ &= g^i_{\rho^-}(O_2^*),\end{aligned}$$

so we have $\ell\Pi(e') \geq \pi(O_2^*)$ and it suffices to bound $\Pi(e')$.

Let L be the unbounded feasible set in P_{ρ^-}; note that its minimum density is ρ^-. Since L is unbounded, there exists j such that $s_j(L) > \bar{\mathsf{s}}_j(e')$. We know that $G_{\rho^-,j}$ maximizes $s_j(\cdot)$ among feasible sets contained in $P_{\rho^-} \setminus \{e_{\rho^-}\}$. Therefore, we have $s_j(G_{\rho^-,j}) + s_j(e') \geq s_j(L) > \bar{\mathsf{s}}_j(e')$.

Now consider two cases. Suppose that for all i, $s_i(G_{\rho^-,i}) \leq \bar{\mathsf{s}}_i(e')$. Then, $s_i(G_{\rho^-,j}) \leq \bar{\mathsf{s}}_i(e')$. For brevity, we denote $G_{\rho^-,j}$ by G'. For all i, let α_i be such that $s_i(G') + \alpha_i s_i(e') = \bar{\mathsf{s}}_i(e')$ and let $\alpha = \min_i \alpha_i$. Without loss of generality, we assume that α_1 is the minimum. By definition, we have $s_1(G' + \alpha e') = \bar{\mathsf{s}}_1(e')$ and $s_i(G' + \alpha e') \leq \bar{\mathsf{s}}_i(e')$ for $i \neq 1$. This implies that $G' + \alpha e'$ is a ρ^--bounded fractional set and so

$$\begin{aligned}\pi(G' + \alpha e') &\geq g^1_{\rho^-}(G' + \alpha e') \\ &= \rho_1^- s_1(G' + \alpha e') - \mathsf{C}_1(s_1(G' + \alpha e')) \\ &= \rho_1(e')\bar{\mathsf{s}}_1(e') - \mathsf{C}_1(\bar{\mathsf{s}}_1(e')) \\ &= \Pi(e').\end{aligned}$$

Since $e' \in \mathbb{I}$, we have $\ell(\pi(G') + \pi(e')) \geq \ell\Pi(e') \geq \pi(O_2^*)$.

In the second case, there exists a dimension i such that $\bar{\mathsf{s}}_i(e') < s_i(G_{\rho^-,i})$. Without loss of generality, we assume that it is the first dimension. In this case, we define G' to be $G_{\rho^-,1}$. Let μ be the minimum density in G'. Note that $\mu_1 > \rho_1^- 1 <$ by definition of the $G_{\rho^-,i}$'s. Since G' is bounded, we have $\bar{\mathsf{s}}_1(e') < s_1(G') \leq \bar{\mathsf{s}}_1(\mu_1)$. Applying Lemma 21 over μ_1 gives

$$g^1_\mu(G') = \mu_1 s_1(G') - \mathsf{C}_1(s_1(G')) > \mu_1 \bar{\mathsf{s}}_1(e') - \mathsf{C}_1(\bar{\mathsf{s}}_1(e')).$$

Since G' is μ-bounded, by Lemma 35 we have

$$\begin{aligned}\pi(G') &\geq g_{\mu}^{1}(G') \\ &> \mu_1 \bar{\mathsf{s}}_1(e') - \mathsf{C}_1(\bar{\mathsf{s}}_1(e')) \\ &> \rho_1^- \bar{\mathsf{s}}_1(e') - \mathsf{C}_1(\bar{\mathsf{s}}_1(e')) \\ &= \Pi(e').\end{aligned}$$

This gives us $\ell\pi(G') = \pi(O_2^*)$. Overall, we get $\ell(\max_i \pi(G_{\rho^-,i}) + \pi(e')) \geq \pi(O_2^*)$. □

We are now ready to give an offline algorithm. Let $D_{\gamma,i}$ be the result of running Algorithm 3 on P_γ with objective π_i and feasibility constraint $\mathcal{F}$.

Algorithm 8 Offline algorithm for multi-dimensional $(\pi, \mathcal{F})$

Let $G_{\max} = \operatorname{argmax}_{G \in \{G_{\gamma,i}\}_{\gamma,i}} \pi(G)$.
Let $D_{\max} = \operatorname{argmax}_{D \in \{D_{\gamma,i}\}_{\gamma,i}} \pi(D)$.
Let $e_{\max} = \operatorname{argmax}_{e \in U} \pi(e)$.
return the most profitable of $G_{\max}$, $D_{\max}$, and $e_{\max}$.

Finally, we use the above lemmas to lower bound the performance of Algorithm 8.

Theorem 37. *Algorithm 8 gives a 7ℓ-approximation to $(\pi, \mathcal{F})$ where π is an ℓ-dimensional profit function and $\mathcal{F}$ is a matroid feasibility constraint.*

Proof. First, we consider only elements from $\mathbb{I}$. If all feasible sets are bounded, then we get $\pi(D_{\max}) \geq \frac{1}{4\ell}\pi(O^*)$ by Lemma 35. On the other hand, if there exists an unbounded feasible set, then we have

$$\begin{aligned}\ell(4\pi(D_{\max}) + \pi(e_{\max}) + \pi(G_{\max})) &\geq 4\sum_i \pi(D_i') + \ell(\pi(e') + \max_i \pi(G_{\rho^-,i})) \\ &\geq \pi(O^* \cap \mathbb{I})\end{aligned}$$

by Lemma 36. For elements from $\mathbb{F}$, Lemma 14 shows that $\pi(e_{\max}) \geq \pi(O^* \cap \mathbb{F})/\ell$. This proves that the algorithm achieves a 7ℓ-approximation. □

E.1 The online setting

In this subsection, we develop the online algorithm for constrained multidimensional profit maximization. First we remark that density prefixes are well-defined in this setting, so we can use Algorithm 4, the online algorithm for single-dimensional constrained profit maximization, in steps 6 and 12. In order to mimic the offline algorithm, we guess (with equal probability) among one of the following scenarios and apply the appropriate subroutine for profit maximization: if there is a single high profit element we apply Dynkin's algorithm, if all feasible sets are bounded we apply Algorithm 4 along a randomly selected dimension, else if there exists an unbounded feasible set we first estimate a density threshold by sampling and then apply Algorithm 4 over a random dimension but restricted to elements with density above the threshold.

Most of this subsection is devoted the analysis of Algorithm 9 when there exists an unbounded feasible set. We prove the overall competitive ratio of the algorithm in Theorem 40.

We define ρ^+ as follows.

Algorithm 9 Online algorithm for multi-dimensional $(\pi, \mathcal{F})$

```
Let c be a uniformly random draw from {1, 2, 3}.
Let i* be a uniformly random draw from {1, ..., ℓ}.
if c = 1 then
    return O_0: the result of running Dynkin's online algorithm.
else if c = 2 then
    return O_1: the result of running Algorithm 4 on (π_{i*}, F).
else if c = 3 then
    Draw k from Binomial(n, 1/2).
    Let the sample X be the first k elements and Y be the remaining elements.
    Determine A(X) using the offline Algorithm 8.
    Let β be a specified constant and let τ be the highest density such that π(A(P_τ^X)) ≥ β/(49ℓ^2) π(A(X)).
    return O_2: the result of running Algorithm 4 on P_τ^Y with objective π_{i*} and feasibility constraint F.
end if
```

Definition 5. *For a fixed parameter $\beta \leq 1$, let ρ^+ be the highest density such that $\pi(O^*(P_{\rho^+})) \geq \left(\frac{\beta}{49\ell^2}\right)^2 \pi(O^*)$.*

The following lemma is essentially an analogue of Lemma 12.

Lemma 38. *Suppose there exists an unbounded feasible set. For fixed parameters $k_1 \geq 1$, $k_4 \leq 1$, $\beta \leq 1$ and $\beta' \leq 1$ assume that there does not exists an element with profit more than $\frac{1}{k_1}\pi(O^*)$. Then with probability at least k_4, we have that τ, as defined in Algorithm 9, satisfies*

1. *$\rho^+ \geq \tau \geq \rho^-$ and*
2. *$\pi(O^*(P_\tau^Y)) \geq \beta' \left(\frac{\beta}{49\ell^2}\right)^2 \pi(O^*)$*

Proof Sketch: We define events E_1 and E_2 as in the single-dimensional setting:

$$E_1 : \pi(O^*(P_{\rho^-}^X)) \geq \beta\pi(O^*(P_{\rho^-}))$$
$$E_2 : \pi(O^*(P_{\rho^+}^Y)) \geq \beta'\pi(O^*(P_{\rho^+})).$$

We observe that the proof of Lemma 12 primarily depends on the properties of density prefixes, in particular that $\pi(O^*(P_{\rho^-}^X))$ is a sufficiently large fraction of $\pi(O^*(P_{\rho^-}))$, and that Algorithm 3's profit does not decrease when considering larger prefixes. Then Theorem 37 gives us $\pi(O^*(P_{\rho^-})) \geq \frac{1}{7\ell}\pi(O^*)$.

Thus, a proof similar to that of Lemma 12 validates Lemma 38. □

The following lemma establishes the competitive ratio of Algorithm 9 in the presence of an unbounded feasible set.

Lemma 39. *Suppose that there exists an unbounded feasible set, and let α be the competitive ratio that Algorithm 4 achieves for a single-dimensional problem $(\pi_i, \mathcal{F})$. Then, for a fixed set Y and threshold τ, satisfying $\tau \geq \rho^-$, we have $\mathrm{E}_\sigma[\pi(O_2)] \geq \frac{1}{\alpha\ell}\,\pi(O^*(P_\tau^Y))$, where the expectation is over all permutations σ of Y.*

Proof. First, we prove that O_2 is bounded. By the definition of ρ^-, since $O_2 \subseteq P_\tau^Y$ it suffices to show that $P_\tau^Y \subseteq \bar{P}_{\rho^-}$, as all feasible sets in $\bar{P}_{\rho^-}$ are bounded. The threshold τ is either equal to or strictly greater than ρ^- and it suffices to consider the former case. If $\tau = \rho^-$, then e_{ρ^-} must have been in the sample set X so it is not in P_τ^Y and so $P_\tau^Y \subseteq \bar{P}_{\rho^-}$.

Hence, by Lemma 34, for all i we have $\pi(O_2) \geq \pi_i(O_2)$. Applying Theorem 13 where the ground set is P_τ^Y gives us $\mathrm{E}_\sigma[\pi_i(O_2) \mid i^* = i] \geq \frac{1}{\alpha}\pi_i(O^*(P_\tau^Y))$. Therefore, we have

$$\begin{aligned}
\mathrm{E}_\sigma[\pi(O_2)] &= \frac{1}{\ell}\sum_i \mathrm{E}_\sigma[\pi(O_2) \mid i^* = i] \\
&\geq \frac{1}{\ell}\sum_i \mathrm{E}_\sigma[\pi_i(O_2) \mid i^* = i] \\
&\geq \frac{1}{\alpha\ell}\sum_i \pi_i(O^*(P_\tau^Y)) \\
&= \frac{1}{\alpha\ell}\pi(O^*(P_\tau^Y)).
\end{aligned}$$

□

We now prove the main result of this section.

Theorem 40. *Let α denote the competitive ratio of Algorithm 4 for the single-dimensional problem $(\pi_i, \mathcal{F})$. Then Algorithm 9 achieves a competitive ratio of $O(\alpha\ell^5)$ for the multi-dimensional problem $(\pi, \mathcal{F})$.*

Proof. If all feasible sets are bounded, we can apply Lemma 34. Then,

$$\begin{aligned}
\mathrm{E}[\pi(O_1)] &\geq \frac{1}{\ell}\sum_i \mathrm{E}[\pi_i(O_1) \mid i^* = i] \\
&\geq \frac{1}{\alpha\ell}\sum_i \pi_i(O^*) \\
&= \frac{1}{\alpha\ell}\pi(O^*).
\end{aligned}$$

Suppose there exists an element with profit more than $\frac{1}{k_1}\pi(O^*)$. Since we apply the standard secretary algorithm with probability $\frac{1}{3}$, and the standard secretary algorithm is e-competitive, in expectation we get a profit of $\frac{1}{3k_1e}$ times the optimal.

Finally, if no element has profit more than $\frac{1}{k_1}\pi(O^*)$ and there exists an unbounded feasible set, using the second inequality of Lemma 38, Lemma 39, and the fact that we output O_2 with probability $\frac{1}{3}$, we get

$$\begin{aligned}
\mathrm{E}[\pi(O_2)] &\geq \frac{1}{3}\mathrm{E}[\pi(O_2) \mid E_1 \wedge E_2]\Pr[E_1 \wedge E_2] \\
&\geq \frac{k_4}{3\alpha\ell} \cdot \mathrm{E}[\pi(O^*(P_\tau^Y)) \mid E_1 \wedge E_2] \\
&\geq \frac{k_4\beta'}{3\alpha\ell}\left(\frac{\beta}{49\ell^2}\right)^2 \pi(O^*).
\end{aligned}$$

Overall, we have that

$$\mathrm{E}[\pi(O)] \geq \min\left\{\frac{k_4\beta'}{3\alpha\ell}\left(\frac{\beta}{49\ell^2}\right)^2, \frac{1}{3k_1e}, \frac{1}{3\alpha\ell}\right\} \cdot \pi(O^*).$$

Since all the involved parameters are fixed constants we get the desired result. □