Optimal Item Pricing in Online Combinatorial Auctions*

José Correa†.[1], Andrés Cristi[2], Andrés Fielbaum[3], Tristan Pollner[4], and S. Matthew Weinberg[5]

[1]Department of Industrial Engineering, Universidad de Chile
[2]Center for Mathematical Modeling, Universidad de Chile
[3]School of Civil Engineering, University of Sydney
[4]Department of Management Science and Engineering, Stanford University
[5]Department of Computer Science, Princeton University

October 20, 2023

Abstract

We consider a fundamental pricing problem in combinatorial auctions. We are given a set of indivisible items and a set of buyers with randomly drawn monotone valuations over subsets of items. A decision-maker sets item prices and then the buyers make sequential purchasing decisions, taking their favorite set among the remaining items. We parametrize an instance by d, the size of the largest set a buyer may want. Our main result asserts that there exist prices such that the expected (over the random valuations) welfare of the allocation they induce is at least a factor $1/(d+1)$ times the expected optimal welfare in hindsight. Moreover, we prove that this bound is tight. Thus, our result not only improves upon the $1/(4d-2)$ bound of Dütting et al., but also settles the approximation that can be achieved by using item prices. The existence of these prices follows from the existence of a fixed point of a related mapping, and therefore, it is non-constructive. However, we show how to compute such a fixed point in polynomial time, even if we only have sample access to the valuation distributions.

We provide additional results for the special case when buyers' valuations are known (but a posted-price mechanism is still desired), and an improved impossibility result for the special case of prophet inequalities for bipartite matching.

*This work was partially supported by NSF CCF-1955205 and by ANID (Chile) through grants ACT210005, FB210005 and FONDECYT 1220054.

†Corresponding author:J. Correa, Department of Industrial Engineering, Universidad de Chile, correa@uchile.cl

1 Introduction

In combinatorial auctions, a set of valuable items is to be allocated among a set of interested agents. Who should get which items in order to maximize the social welfare? This is a fundamental economic question, and a ubiquitous allocation mechanism is to simply set a price for each item and let the agents buy their preferred subset of items under those prices. The study of these mechanisms dates back to the investigations of Leon Walras over a century ago, and is closely related to the notion of Walrasrian equilibrium. Understanding the existence and approximation of Walrasrian equilibrium and related notions under pricing mechanisms has been an active area of research in recent years [BLNPL14, FGL14, FGL16, BK19, LW20].

In this paper, we follow the approach of online combinatorial auctions and study the welfare achieved by posted-price mechanisms in a very general setup. Specifically, our mechanisms post a price p_i on each item i. Then, buyers with randomly-drawn arbitrary monotone valuations over the subsets of items arrive in arbitrary order, and upon arrival pick their preferred subset among those items that are left (at the posted prices). Of course, in this generality little can be said about the social welfare induced by posted-price mechanisms, so it is common to parametrize the instances by d, the largest size of a set a buyer might be interested in. This parametrization is interesting from a combinatorial perspective: finding a socially optimal allocation is NP-hard already when $d \geq 3$, and even hard to approximate [Tre01]. Moreover, if we restrict the buyers' valuations to be deterministic and single-minded,[1] we recover the classic hypergraph matching problem.

A natural goal in this context is thus to obtain prices for each item, such that the expected welfare of the resulting allocation when adversarial-order buyers iteratively purchase their preferred set is high with respect to the expected welfare of an optimal allocation. In other words, our goal in this paper is to find posted-price mechanisms with good approximation guarantees of the expected welfare of an optimal allocation, as a function of d.

1.1 Our Results

Our main result in this paper is to determine the tight approximation guarantee of item pricing as a function of d. Specifically, we prove that there always exists a posted-price mechanism such that the expected welfare of the resulting allocation when adversarial-order buyers iteratively purchase their preferred set (at the posted prices) is at least a $1/(d+1)$ fraction of the expected welfare of an optimal allocation (Theorem 3.1). In particular, our results improve the bound of $1/(4d-2)$ given in [DFKL20] to $1/(d+1)$, which is tight (Proposition 3.5).

In Section 6, we further consider the special case that arises when valuations are deterministic and buyers are single-minded. In this situation the welfare optimization problem corresponds to matching in a hypergraph with edges of size at most d. So the problem of finding item prices boils down to finding a set of thresholds, one for each vertex, such that the value of the solution in which hyperedges arrive sequentially (in any order) and are greedily included in the solution when their weight is higher that the sum of the corresponding vertex thresholds, is as close as possible to the optimal solution. For the case of standard matching ($d = 2$) we prove that there exist prices guaranteeing a factor of $1/2$ of the optimal solution and that there do not exist prices guaranteeing a factor better than $2/3$. The tight factor is left as an open problem. More generally, we prove that there are prices obtaining a fraction $1/d$ of the optimal solution (thus slightly improving our general $1/(d+1)$), and that it is not possible to do better than $\approx 1/\sqrt{d}$.

Finally, in Section 7, we prove that no bipartite matching prophet inequality can beat a $3/7$ approximation.[2] Note that prophet inequality for matching corresponds to a special case of our main

[1] Note that this is a particular case of monotone valuations. Indeed, a single-minded buyer can be modeled through a fixed set T, such that the buyer values a set S at a certain positive amount if $T \subseteq S$, and at 0 otherwise.

[2] Recall that a prophet inequality instance specifies a set system $\mathcal{S}$ of elements and $\mathcal{I} \subseteq 2^{\mathcal{S}}$ of feasible sets, together with a distribution $\mathcal{D}_e$ for each $e \in S$. The elements of $\mathcal{S}$ are revealed one at a time, together with a draw $X_e \sim \mathcal{D}_e$ independently, at which time the gambler must immediately and irrevocably decide whether to accept e (gaining

result with $d = 2$. Therefore our Proposition 3.5 establishes that no non-adaptive vertex-pricing prophet inequality can guarantee better than a $1/3$-approximation to the expected offline optimum. However, adaptive non-pricing prophet inequalities are known to achieve $.337 > 1/3$ [EFGT20]. So the content of Theorem 7.1 is that no prophet inequality can beat $3/7$, even those that are adaptive and not based on vertex-pricing. It remains an open question to nail the precise approximation guarantee of prophet inequalities for bipartite matchings within $[.337, 3/7]$.

1.2 Context and Related Work

Posted-price Mechanisms. Posted-price mechanisms are ubiquitous within economics and computation owing to their simplicity. They are commonly used as subroutines in truthful mechanisms that approximately maximize welfare [DNS05, KV12, Dob16, AS19, AKS21]. They are also used as subroutines in simple mechanisms to approximately maximize revenue in Bayesian settings [CHMS10, KW12, CM16, CZ17]. Our work considers the same model initiated by [FGL14] (welfare maximization in Bayesian settings). Other works consider restrictions on the valuations, such as subadditive [DKL20], while others consider the unrestricted case [DFKL20]. In this last work [DFKL20] consider unrestricted valuations over sets of size at most d, and derive a collection of prices leading to an approximation guarantee of $1/(4d-2)$. Our paper contributes to this line of work by providing the tight approximation guarantee of $1/(d+1)$ for posted-price mechanisms in this model.

Prophet Inequalities. When there is a single item (and thus $d = 1$) our problem is equivalent to the single-item prophet inequality and thus our result takes the same form as the classic result of [SC84], who proved that the optimal prophet inequality (whose factor is $1/2$) can be achieved with a single threshold. A special case of our problem when buyers are single-minded corresponds to various multiple-choice prophet inequality settings, and our results improve upon the state-of-the-art. In particular, all prophet inequalities deduced from our main result are *non-adaptive*: for each element e, a threshold T_e is set at the beginning of the algorithm. Element e is accepted if and only if the buyers' valuation w_e exceeds the threshold, i.e., $w_e \geq T_e$ (and it is feasible to accept e).

When $d = 2$ and buyers are single-minded, our problem translates into the matching prophet inequality problem. In this setting [GW19] obtained $1/3$-approximation for the case of bipartite graphs. So our paper also contributes by extending this result to general graphs. Note that recent work of [EFGT20] provides a .337-approximation in this case, although it sets thresholds adaptively. We further contribute to the $d = 2$ case by proving that no prophet inequality (adaptive or not) can guarantee better than a $3/7$-approximation for the bipartite matching prophet inequality (Theorem 7.1).

For arbitrary d when buyers are single-minded, our problem translates into the d-dimensional hypergraph prophet inequality, which generalizes the prophet inequality problem over the intersection of d partition matroids. Here, a $1/(4d-2)$-approximation was first given in [KW12], and improved to $1/(e(d+1))$ in [FSZ16]. Our work improves this to $1/(d+1)$, and with non-adaptive thresholds. A lower bound of [KW12, SVW23] proves that it is not possible to achieve an $\omega(1/d^{1/2+1/\log\log d})$ approximation even for this special case, but it remains an open problem to determine the tight ratio for prophet inequalities for the intersection of d partition matroids (and for the d-dimensional hypergraph prophet inequality).

Very recent work establishes a constant-factor prophet inequality for combinatorial auctions with subadditive valuations [CC23]. An important open question at the intersection of this paper and ours is whether their guarantee can also be achieved by a posted-price mechanism.

reward X_e), or reject forever. The gambler must maintain at all times that the collection of accepted elements is in $\mathcal{I}$, and aims to maximize its expected reward. Bipartite matching refers to the case where each element of $\mathcal{S}$ is an edge in a bipartite graph, and $\mathcal{I}$ contains all matchings.

1.3 Brief Technical Highlight

The proof of our main result breaks down the expected welfare into the "revenue" and "utility" achieved by setting prices and searches for properly "balanced thresholds." This idea has been widely used in the context of combinatorial prophet inequalities [KW12, FGL14, GW19, DFKL20]. In particular, we target prices that are "low enough" so that a buyer with high value for some set will choose to purchase it, yet also "high enough" so that the revenue gained when a bidder purchases items they should not receive in the optimal allocation compensates for the lost welfare. In comparison to prior work using this framework, the conditions that guarantee such prices are more involved, and we prove their existence using Brouwer's fixed point theorem.

As our proof makes use of Brouwer's fixed point theorem, it is inherently non-constructive. We however show in Sections 4 and 5 how to compute our prices in polynomial time. In particular, our approach makes use of a configuration LP relaxation to cope with the APX-hardness of optimizing welfare, and a convex optimization formulation to find our fixed point.

1.4 Summary and Roadmap

We precisely define our model in Section 2. Section 3 presents our main result: a posted-price mechanism that achieves a $1/(d+1)$-approximation to the optimal expected welfare, when buyers have arbitrary monotone valuations and are interested in sets of size at most d. Recall that this approximation guarantee is tight (we provide a simple example witnessing this in Proposition 3.5). Section 4 extends some of the bounds derived in Section 3 by using as benchmark the configuration LP. In Section 5, we show how to compute our desired prices in polynomial time. In Section 6, we consider the special case where the distributions are point-masses, and in Section 7, we give an improved impossibility result for bipartite matching prophet inequalities.

2 Model

In our basic model, we have a (multi)set of items M in which there are $k_j \geq 1$ copies of each item $j \in M$.[3] The set of buyers, denoted by N, arrive sequentially (in arbitrary order) and buy some of those items. Each buyer $i \in N$ has a valuation function $v_i : 2^M \to \mathbb{R}_{\geq 0}$, which is randomly and independently chosen according to a given distribution $\mathcal{F}_i$ (defined over a set of possible valuation functions). As it is standard, we assume that each possible realization of each v_i is monotone (i.e., $A \subseteq B \Rightarrow v_i(A) \leq v_i(B)$). We parametrize an instance of the problem by d, the size of the largest set a buyer might be interested in. Thus, if $A \subseteq M$ is such that $|A| > d$, then

$$v(A) = \max_{B \subseteq A, |B|=d} v(B) \tag{1}$$

Note that while there are $k_j \geq 1$ copies of each item $j \in M$, no bidder achieves value from additional copies (and therefore, without loss of generality, bidders cannot buy more than one copy).

In this paper, we are interested in exploring the limits of using item prices as a mechanism to assign items to buyers. In a pricing mechanism, we set item prices $p \in \mathbb{R}^M_{\geq 0}$ and then consider an arbitrary arrival order of the buyers (note different copies of the same item must have the same price). Thus, buyer i buys the set of remaining items according to

$$\max_{A \subseteq R_i} \left(v_i(A) - \sum_{j \in A} p_j \right), \tag{2}$$

where R_i denotes the items for which there remains an unsold copy when i arrives. Note that Eq. (2) might be solved by $A = \emptyset$, i.e., buyer i might opt not to buy anything. When there is a tie

[3] Throughout the paper M is actually a set and refers to the set of different items.

between different sets, the buyer can choose arbitrarily, implying that our results need to be valid even for the worst-case tiebreaking.[4]

More precisely, if σ is the arrival order of the buyers, so that buyer i comes at time $\sigma(i)$, then buyer i gets the set $B_i(\sigma) = \arg\max_{A \subseteq R_i(\sigma)} \left(v_i(A) - \sum_{j \in A} p_j \right)$, where $R_i(\sigma) = \{j \in M : k_j > |\{\ell \in N : \sigma(\ell) < \sigma(i) \text{ and } j \in B_\ell(\sigma)\}|\}$. With this, given an instance of the problem (determined by M, k_j for all $j \in M$, N, and $\mathcal{F}_i$ for all $i \in N$), the quality measure of a price vector $p \in \mathbb{R}^M_{\geq 0}$ is the worst case (over the arrival orders) expected (over the valuations) welfare of the allocation it induces. Denoting this quantity by $ALG(p)$ we have that:

$$ALG(p) := \min_\sigma \mathbb{E} \left(\sum_{i \in N} v_i(B_i(\sigma)) \right).$$

On the other hand, the benchmark we compare to throughout the paper is the expected value of the optimal welfare-maximizing allocation, OPT, formally defined as

$$OPT := \mathbb{E} \left(\max_{\{A_i\}_{i \in N}} \left\{ \sum_{i \in N} v_i(A_i) : \text{ s.t. } |\{i \in N : j \in A_i\}| \leq k_j, \text{ for all } j \in M \right\} \right).$$

We denote by OPT_i the random set that buyer i gets in an optimal allocation.

In Section 6 we consider the special case of our problem in which

(i) valuations are deterministic,

(ii) there is a single copy of each item (i.e., $k_j = 1$ for all $j \in M$), and

(iii) buyers are *single-minded*, i.e., each buyer i has a set A_i, with $|A_i| \leq d$, such that $A_i \not\subseteq B \Rightarrow v_i(B) = 0, A_i \subseteq B \Rightarrow v_i(B) = v_i(A_i)$.

Interestingly, already in this particular setup, the problem of maximizing the welfare of an allocation corresponds to the classic NP-hard combinatorial optimization problem of hypergraph matching with hyperedges of size at most d. Indeed, in an optimal allocation buyer i either gets A_i or $\emptyset$, implying that maximizing the (now deterministic) welfare of the allocation is equivalent to finding a subset of pairwise disjoint A_i's of maximum total valuation.

In Section 7 we continue to focus on single-minded buyers with one copy of every item, but allow random valuations; this captures the *prophet inequality* model with a large class of feasibility constraints (we focus on bipartite matching).

3 Main Result: A $1/(d+1)$-approximation for Random Valuations

In this section we prove there exists a vector of item prices such that the resulting allocation yields in expectation at least a $1/(d+1)$ fraction of the optimal social welfare. Additionally, we show that this bound is tight.

Theorem 3.1. *There exists a vector of prices $p \in \mathbb{R}^M_{\geq 0}$ such that*

$$(d+1) \cdot ALG(p) \geq OPT.$$

To prove the theorem we will make use of the following function. For each $A \subseteq M$ and $i \in N$, we define $z_{i,A} : \mathbb{R}^M_{\geq 0} \to \mathbb{R}$ as

$$z_{i,A}(p) := \mathbb{E} \left(\mathbb{1}_{OPT_i = A} \cdot \left[v_i(A) - \sum_{j \in A} p_j \right]_+ \right),$$

[4] In some of the constructions in Section 6 we break ties conveniently but all the results hold by slightly tweaking the instances.

where $[x]_+$ denotes $\max\{x, 0\}$. The function $z_{i,A}(p)$ can be interpreted as follows: imagine we calculate the optimal allocation and offer buyer i the set OPT_i at the prices given by p. Then, $z_{i,A}$ would be the contribution of the set A to the non-negative part of the expected utility of buyer i.

We assume without loss of generality that $|OPT_i| \leq d$ for all $i \in N$, so $z_{i,A}(p) = 0$ if $|A| > d$. We start by showing a lower bound for $ALG(p)$ in terms of the values $z_{i,A}(p)$. This type of analysis is, by now, standard in combinatorial prophet inequalities [FGL14, KW12, GW19].

Lemma 3.2. *For any vector of prices $p \in \mathbb{R}^M_{\geq 0}$,*

$$ALG(p) \geq \min_{C \subseteq M} \left\{ \sum_{j \notin C} k_j \cdot p_j + \sum_{i \in N} \sum_{A \subseteq C} z_{i,A}(p) \right\}.$$

Proof. In this proof we assume the arrival order σ is arbitrary, and for simplicity we denote $B_i(\sigma)$ and $R_i(\sigma)$ simply by B_i and R_i. We separate the welfare of the resulting allocation into *revenue* and *utility*, i.e., we separate $\sum_{i \in N} v_i(B_i)$ into

$$\text{Revenue} = \sum_{i \in N} \sum_{j \in B_i} p_j \quad \text{and} \quad \text{Utility} = \sum_{i \in N} \left(v_i(B_i) - \sum_{j \in B_i} p_j \right).$$

Recall that R_i is the set of items with remaining copies when i arrives. Similarly, denote by R the set of items that have remaining copies by the end of the process. Note first that

$$\mathbb{E}(\text{Revenue}) \geq \mathbb{E}\left(\sum_{j \notin R} k_j \cdot p_j \right).$$

This is simply because each item $j \notin R$ has had all k_j copies purchased. As for the utility, for any $i \in N$, by the definition of B_i it holds that

$$v_i(B_i) - \sum_{j \in B_i} p_j = \max_{A \subseteq R_i} v_i(A) - \sum_{j \in A} p_j$$

Note now that v_i and R_i are independent. Let $(\tilde{v}_i)_{i \in N}$ be independent realizations of the valuations, and $\widetilde{OPT}_i$ the corresponding optimal solution. With this, and noting that $R \subseteq R_i$, we can rewrite the expected utility of agent i as

$$\mathbb{E}\left(\max_{A \subseteq R_i} v_i(A) - \sum_{j \in A} p_j \right) = \mathbb{E}\left(\max_{A \subseteq R_i} \tilde{v}_i(A) - \sum_{j \in A} p_j \right) \geq \mathbb{E}\left(\max_{A \subseteq R} \tilde{v}_i(A) - \sum_{j \in A} p_j \right). \tag{3}$$

We replace the maximization over subsets of R with a particular choice, $\widetilde{OPT}_i$, whenever it is contained in R and gives positive utility (otherwise we take $\emptyset$). This obtains the following lower bound on the expected utility of agent i:

$$\begin{aligned}
\mathbb{E}\left(\mathbb{1}_{\{\widetilde{OPT}_i \subseteq R\}} \cdot \left[\tilde{v}_i(\widetilde{OPT}_i) - \sum_{j \in \widetilde{OPT}_i} p_j \right]_+ \right) &= \mathbb{E}\left(\sum_{A \subseteq R} \mathbb{1}_{\{\widetilde{OPT}_i = A\}} \cdot \left[\tilde{v}_i(A) - \sum_{j \in A} p_j \right]_+ \right) \\
&= \mathbb{E}\left(\sum_{A \subseteq R} \mathbb{E}\left(\mathbb{1}_{\{\widetilde{OPT}_i = A\}} \cdot \left[\tilde{v}_i(A) - \sum_{j \in A} p_j \right]_+ \right) \right) \\
&= \mathbb{E}\left(\sum_{A \subseteq R} z_{i,A}(p) \right). \qquad (4)
\end{aligned}$$

Summing over all agents, we get that

$$\mathbb{E}(\text{Utility}) \geq \mathbb{E}\left(\sum_{i\in N}\sum_{A\subseteq R} z_{i,A}(p)\right).$$

Therefore, adding the revenue and the utility we get that

$$ALG(p) \geq \mathbb{E}\left(\sum_{j\notin R} k_j \cdot p_j + \sum_{i\in N}\sum_{A\subseteq R} z_{i,A}(p)\right).$$

Replacing the expectation over R with a minimization over subsets of M yields the bound of the lemma. □

Lemma 3.3. *For any vector of prices $p \in \mathbb{R}^M_{\geq 0}$,*

$$OPT \leq \sum_{j\in M} k_j \cdot p_j + \sum_{i\in N}\sum_{A\subseteq M} z_{i,A}(p).$$

Proof. We have that OPT equals

$$\sum_{i\in N}\mathbb{E}(v_i(OPT_i)) = \mathbb{E}\left(\sum_{i\in N}\sum_{j\in OPT_i} p_j\right) + \sum_{i\in N}\mathbb{E}\left(v_i(OPT_i) - \sum_{j\in OPT_i} p_j\right).$$

Now we upper bound these two terms separately. Note that in the first term each item $j \in M$ appears at most k_j times, so

$$\mathbb{E}\left(\sum_{i\in N}\sum_{j\in OPT_i} p_j\right) \leq \sum_{j\in M} k_j \cdot p_j.$$

For the second part, we upper bound with the positive part of the difference, and sum over all possible realizations of OPT_i:

$$\begin{aligned}\sum_{i\in N}\mathbb{E}\left(v_i(OPT_i) - \sum_{j\in OPT_i} p_j\right) &= \sum_{i\in N}\sum_{A\subseteq M}\mathbb{E}\left(\mathbb{1}_{\{OPT_i=A\}}\left(v_i(A) - \sum_{j\in A} p_j\right)\right)\\ &\leq \sum_{i\in N}\sum_{A\subseteq M} z_{i,A}(p).\end{aligned}$$

Putting together the two upper bounds we obtain the bound on OPT. □

Lemmas 3.2 and 3.3 provide a similar form to lower bound $ALG(p)$ and upper bound OPT as a function of p. Now, we will prove the existence of a good choice of p where these bounds differ by at most a factor of $d+1$.

Lemma 3.4. *There exists a vector of prices $p \in \mathbb{R}^M_{\geq 0}$ such that for every $j \in M$ we have*

$$p_j = \frac{1}{k_j}\sum_{i\in N}\sum_{A\subseteq M: j\in A} z_{i,A}(p).$$

Proof. The proof will be an application of Brouwer's fixed point theorem. Let K denote the compact set $K := [0, OPT]^M \subseteq \mathbb{R}^M_{\geq 0}$. We define a function $\psi : K \to K$ as follows: for a vector $p \in K$ and item $j \in M$, the j^{th} coordinate of ψ is

$$\psi_j(p) = \frac{1}{k_j}\sum_{i\in N}\sum_{A\subseteq M: j\in A} z_{i,A}(p). \tag{5}$$

We prove now that ψ is a well-defined continuous function, from the compact set K into itself, and therefore it has a fixed point by Brouwer's fixed point theorem. Note that a fixed point of ψ is exactly the vector of prices we are looking for.

In fact, recall that we defined $z_{i,A}(p) = \mathbb{E}(\mathbb{1}_{OPT_i=A} \cdot [v_i(A) - \sum_{j\in A} p_j]_+)$, which is a nonincreasing function of p_j, for all $j \in M$. Moreover, note that since $[\cdot]_+$ is a convex function, $z_{i,A}$ is also a convex function of p_j for all $j \in M$. The monotonicity of $z_{i,A}$ implies that for all $p \in K$ and $j \in M$, $\psi_j(p) \le \psi_j(0) \le \frac{1}{k_j} OPT$, and therefore $\psi(p) \in K$ for all $p \in K$. The convexity of $z_{i,A}$ implies it is also continuous, so ψ is a continuous function.

We've now argued that ψ is a continuous function from K to itself, and therefore a fixed point exists, which proves the lemma. $\square$

Proof of Theorem 3.1. Using the vector of prices from Lemma 3.4, we apply the bound of Lemma 3.2 and conclude

$$\begin{aligned}
ALG(p) &\ge \min_{C\subseteq M} \left\{ \sum_{j\notin C} k_j \cdot p_j + \sum_{i\in N} \sum_{A\subseteq C} z_{i,A}(p) \right\}. \\
&\ge \min_{C\subseteq M} \left\{ \sum_{j\notin C} k_j \cdot \frac{1}{k_j} \sum_{i\in N} \sum_{A\subseteq M: j\in A} z_{i,A}(p) + \sum_{i\in N} \sum_{A\subseteq C} z_{i,A}(p) \right\} \\
&= \min_{C\subseteq M} \left\{ \sum_{i\in N} \sum_{A\subseteq M} z_{i,A}(p) \cdot (|A \setminus C| + \mathbb{1}_{A\subseteq C}) \right\} \\
&\ge \min_{C\subseteq M} \left\{ \sum_{i\in N} \sum_{A\subseteq M} z_{i,A}(p) \cdot (\mathbb{1}_{|A\setminus C|\ge 1} + \mathbb{1}_{A\subseteq C}) \right\} \\
&= \sum_{i\in N} \sum_{A\subseteq M} z_{i,A}(p).
\end{aligned}$$

For OPT, substituting our fixed point in the upper bound of Lemma 3.3 gives

$$\begin{aligned}
OPT &\le \sum_{j\in M} \sum_{i\in N} \sum_{A\subseteq M: j\in A} z_{i,A}(p) + \sum_{i\in N} \sum_{A\subseteq M} z_{i,A}(p) \\
&= \sum_{i\in N} \sum_{A\subseteq M} (|A| + 1) \cdot z_{i,A}(p) \\
&\le (d+1) \sum_{i\in N} \sum_{A\subseteq M} z_{i,A}(p).
\end{aligned}$$

Comparing the two bounds we see that $(d+1) \cdot ALG(p) \ge OPT$. $\square$

To wrap up the section, we establish that the bound of Theorem 3.1 is best possible by modifying a simple example of [FGL14]. This example establishes a lower bound of d if we restrict to deterministic valuations. Here we add stochasticity to match the bound of Theorem 3.1.

Before doing so, we note that even without stochasticity, the bound of $d+1$ is tight for our fixed point approach. For this, consider a deterministic instance with d items and two single-minded buyers: buyer A has value $v = 1/(d+1) < 1$ for a single item and buyer B has value 1 for the set of all items. Then the fixed point sets price $1/(d+1)$ on all items, since the fixed point equations are $p_j = 1 - \sum_{j\in M} p_j$, for all $j \in M$ (Buyer B has utility $1/(d+1)$ at these prices). Then the online algorithm will assign an item to buyer A so that the bound of $d+1$ is tight.

Proposition 3.5. *For all d, and all $\delta > 0$, there exists an instance on $|N| = 2$ buyers and $|M| = d$ items such that for all p, $ALG(p) \le 1$, yet $OPT = d + 1 - \delta$.*

Proof. Consider a set M of exactly d items with a single copy of each, and a very small $\varepsilon > 0$. There are two buyers. Buyer A values any nonempty subset of the items at 1. Buyer B only assigns value to getting all d items, and this value is $d - \varepsilon$ with probability $1 - \varepsilon$ and it is $1/\varepsilon$ with probability ε.

In any instance where buyer A purchases a non-empty subset, the resulting social welfare is 1. Note that this is certain to happen if we set the prices so that $\sum_{j \in M} p_j < d$ and buyer A arrives before buyer B. If, on the contrary, $1/\varepsilon \geq \sum_{j \in M} p_j \geq d$ and buyer A does not purchase anything, buyer B will only purchase items with probability ε. In this case, the expected total welfare is also 1. This establishes that $ALG(p) = 1$ for all p such that $\sum_{j \in M} p_j \leq 1/\varepsilon$, and $ALG(p) = 0$ otherwise. Finally, it is clear that in this instance the optimal welfare is achieved by always assigning all items to buyer B, which results in an expected welfare of $(d - \varepsilon) \cdot (1 - \varepsilon) + \varepsilon \cdot (1/\varepsilon) \geq d + 1 - (d+1)\varepsilon$. Setting $\varepsilon = \delta/(d+1)$ completes the proof. □

4 Bounds using an optimal LP solution

Our next step is to make Theorem 3.1 constructive and show that the underlying prices can be efficiently computed. To this end, we need to establish that Lemmas 3.2 and 3.3 also hold when in the definition of $z_{i,A}(p)$ we replace $\mathbb{1}_{OPT_i = A}$ with $x_{i,A}$, an optimal solution of the linear relaxation of the optimal allocation problem. This means we replace $z_{i,A}(p)$ with

$$\tilde{z}_{i,A}(p) = \mathbb{E}\left(x_{i,A} \cdot \left[v_i(A) - \sum_{j \in A} p_j \right]_+ \right),$$

where $x = (x_{i,A})_{i \in N, A \subseteq M}$, solves

$$\begin{aligned} \text{(LP)} \qquad & \max_{x \geq 0} \sum_{i \in N} \sum_{A \subseteq M} x_{i,A} \cdot v_i(A) \\ & \text{s.t.} \sum_{A \subseteq M} x_{i,A} \leq 1, \text{ for all } i \in N, \\ & \sum_{i \in N} \sum_{A: j \in A} x_{i,A} \leq k_j, \text{ for all } j \in M. \end{aligned}$$

These results will be used in the next section and their proofs are almost identical to the original ones in Section 3.

Lemma 4.1. *For any vector of prices $p \in \mathbb{R}^M_{\geq 0}$,*

$$ALG(p) \geq \min_{C \subseteq M} \left\{ \sum_{j \notin C} k_j \cdot p_j + \sum_{i \in N} \sum_{A \subseteq C} \tilde{z}_{i,A}(p) \right\}.$$

Proof. In this proof we assume the arrival order σ is arbitrary, and for simplicity we denote $B_i(\sigma)$ and $R_i(\sigma)$ simply by B_i and R_i. We separate the welfare of the resulting allocation into *revenue* and *utility*, i.e., we separate $\sum_{i \in N} v_i(B_i)$ into

$$\text{Revenue} = \sum_{i \in N} \sum_{j \in B_i} p_j \quad \text{and} \quad \text{Utility} = \sum_{i \in N} \left(v_i(B_i) - \sum_{j \in B_i} p_j \right).$$

Recall that R_i is the set of items with remaining copies when i arrives. Similarly, denote by R the set of items that have remaining copies by the end of the process. Note first that

$$\mathbb{E}(\text{Revenue}) \geq \mathbb{E}\left(\sum_{j \notin R} k_j \cdot p_j \right).$$

This is simply because all items $j \notin R$ have sold all k_j copies. For the utility, for any $i \in N$, by the definition of B_i it holds that

$$v_i(B_i) - \sum_{j \in B_i} p_j = \max_{A \subseteq R_i} v_i(A) - \sum_{j \in A} p_j$$

Note now that v_i and R_i are independent. Let $(\tilde{v}_i)_{i \in N}$ be independent realizations of the valuations. With this and noting that $R \subseteq R_i$, we can rewrite the expected utility of agent i as

$$\mathbb{E}\left(\max_{A \subseteq R_i} v_i(A) - \sum_{j \in A} p_j\right) = \mathbb{E}\left(\max_{A \subseteq R_i} \tilde{v}_i(A) - \sum_{j \in A} p_j\right) \geq \mathbb{E}\left(\max_{A \subseteq R} \tilde{v}_i(A) - \sum_{j \in A} p_j\right). \tag{6}$$

Let $\tilde{x}$ denote an optimal solution of LP when the values are $(\tilde{v}_i)_{i \in N}$. Since $(\tilde{v}_i)_{i \in N}$ is independent of R, $\tilde{x}$ is also independent of R. Since $\tilde{x}$ is feasible for LP, for any given $i \in N$, $\sum_{A \subseteq R} \tilde{x}_{i,A} \leq \sum_{A \subseteq M} \tilde{x}_{i,A} \leq 1$. We can replace the maximization over subsets of R in Eq. (6) with the convex combination of particular choices given by $(\tilde{x}_{i,A})_{A \subseteq R}$. Thus, we obtain the following lower bound.

$$\begin{aligned}
\mathbb{E}\left(\max_{A \subseteq R} \tilde{v}_i(A) - \sum_{j \in A} p_j\right) &\geq \mathbb{E}\left(\sum_{A \subseteq R} \tilde{x}_{i,A} \cdot \left[\tilde{v}_i(A) - \sum_{j \in A} p_j\right]_+\right) \\
&= \mathbb{E}\left(\sum_{A \subseteq R} \mathbb{E}\left(\tilde{x}_{i,A} \cdot \left[\tilde{v}_i(A) - \sum_{j \in A} p_j\right]_+\right)\right) \\
&= \mathbb{E}\left(\sum_{A \subseteq R} \tilde{z}_{i,A}(p)\right).
\end{aligned} \tag{7}$$

The positive part $[\cdot]_+$ comes from the fact that we can always choose $\emptyset \subseteq R$ in the maximization in Eq. (6). Summing over all agents, we get that

$$\mathbb{E}(\text{Utility}) \geq \mathbb{E}\left(\sum_{i \in N} \sum_{A \subseteq R} \tilde{z}_{i,A}(p)\right).$$

Therefore, adding the revenue and the utility we get that

$$ALG(p) \geq \mathbb{E}\left(\sum_{j \notin R} k_j \cdot p_j + \sum_{i \in N} \sum_{A \subseteq R} \tilde{z}_{i,A}(p)\right).$$

Replacing the expectation over R with a minimization over subsets of M we obtain the bound of the lemma. ☐

Lemma 4.2. *For any vector of prices $p \in \mathbb{R}^M_{\geq 0}$,*

$$OPT \leq \sum_{j \in M} k_j \cdot p_j + \sum_{i \in N} \sum_{A \subseteq M} \tilde{z}_{i,A}(p).$$

Proof. Let x be an optimal solution of LP. We have that

$$\begin{aligned}
OPT &\leq \mathbb{E}\left(\sum_{i \in N} \sum_{A \subseteq M} x_{i,A} \cdot v_i(A)\right) \\
&= \mathbb{E}\left(\sum_{i \in N} \sum_{A \subseteq M} x_{i,A} \cdot \sum_{j \in A} p_j\right) + \mathbb{E}\left(\sum_{i \in N} \sum_{A \subseteq M} x_{i,A} \cdot \left(v_i(A) - \sum_{j \in A} p_j\right)\right).
\end{aligned}$$

Now we upper bound these two terms separately. Since x is feasible for LP, for all $j \in M$ we have that $\sum_{i\in N}\sum_{A:j\in A} x_{i,A} \le k_j$, so the first term satisfies

$$\mathbb{E}\left(\sum_{i\in N}\sum_{A\subseteq M} x_{i,A}\cdot\sum_{j\in A} p_j\right) \le \sum_{j\in M} k_j\cdot p_j.$$

For the second term we simply upper bound the difference with its positive part.

$$\begin{aligned}
&\mathbb{E}\left(\sum_{i\in N}\sum_{A\subseteq M} x_{i,A}\cdot\left(v_i(A)-\sum_{j\in A} p_j\right)\right)\\
&\le \sum_{i\in N}\sum_{A\subseteq M}\mathbb{E}\left(x_{i,A}\cdot\left[v_i(A)-\sum_{j\in A} p_j\right]_+\right)\\
&\le \sum_{i\in N}\sum_{A\subseteq M}\tilde{z}_{i,A}(p).
\end{aligned}$$

Putting together the two upper bounds we obtain the bound on OPT. □

5 Efficient Computation

So far, our main result is nonconstructive for several reasons. First, it requires a fixed-point computation (which is PPAD-hard in general). Second, evaluating the function for which we hope to find a fixed point requires computing the expected value of a random variable with exponential support (which is #P-hard in general). Finally, sampling the random variable whose expected value defines our function requires computing the optimal allocation, which is NP-hard in general, even to approximate. As a first step, in the previous section, we showed that we can replace the optimal solution with the solution to a (exponentially large) linear program.

In this section, we show how to overcome all three barriers, and efficiently (in time polynomial in $|M|$ and $|N|$) compute the prices, even when d is not a constant. Notice that when d is not a constant, a complete description of the distributions, or even of a single deterministic valuation function, might be exponentially large. Thus, we assume instead that we can draw samples from the distributions of valuation functions, which we access in a black-box manner via *demand queries*.

Definition 5.1. A demand query of a valuation function $v : 2^M \to \mathbb{R}_{\ge 0}$ accepts a price vector $p \in \mathbb{R}^M_{\ge 0}$ and returns a subset of items A such that $v(A) = \max_{B\subseteq M}\left(v(B) - \sum_{j\in B} p_j\right)$.

We note that while a demand query only returns the subset A and not the associated valuation $v(A)$, we can compute the valuation of any subset using polynomially-many demand queries [NRTV07, Lemma 11.22].

Theorem 5.1. *If there is a number $v_{\max}$ such that $v_i(A) \le v_{\max}$ for all $i \in N, A \subseteq M$ with probability 1, and such that $v_{\max}/OPT \le \mathrm{poly}(|M|,|N|)$, then for every $\varepsilon > 0$ we can calculate prices $\hat{p}$ such that*

$$(d+1+\varepsilon)\cdot ALG(\hat{p}) \ge OPT,$$

with probability $1-\varepsilon$, in time $\mathrm{poly}(|M|,|N|,1/\varepsilon)$, using $\mathrm{poly}(|M|,|N|,1/\varepsilon)$ samples of the valuations and $\mathrm{poly}(|M|,|N|,1/\varepsilon)$ demand queries in total.

To prove this theorem we first show that under the condition $v_{\max}/OPT \le \mathrm{poly}(|M|,|N|)$ we can approximate the function ψ (as defined in Section 3) using polynomially-many samples. This approximation requires computing the optimal fractional allocation, which can be done in polynomial time using demand queries. Finally, we show that the structure of ψ allows us to efficiently compute a fixed point through a convex quadratic program.

Even though $v_{\max}/OPT \leq \mathrm{poly}(|M|, |N|)$ is a seemingly strong condition, the following example illustrates its necessity in our approach. Consider an instance with one item and two buyers. For a small $\delta > 0$, the first buyer has a valuation of δ for the item, while the second has a valuation of $1/\delta$ with probability δ, and 0 otherwise. In this instance $OPT = 1 + \delta - \delta^2$. Most of the time the optimal allocation gives the item to the first buyer; however, most of the value in OPT comes from the second buyer. Thus, in order to obtain a good approximation of OPT using samples, we need to sample the valuation functions enough times to see at least once the $1/\delta$ valuation of the second buyer, i.e., we require $\Omega(1/\delta)$ samples. Otherwise, we could not distinguish the instance from one where the second buyer has valuation identically 0 (in which case we should allocate the item to the first buyer, obtaining a welfare of only δ).

5.1 Proof of Theorem 5.1

Our strategy to find the prices has several steps. First, we use an estimate $\hat{\psi}$ of ψ (recall the definition of ψ from Section 3). The function $\hat{\psi}$ differs from ψ in two ways: first, it replaces the optimal integral allocation with the optimal fractional allocation in the definition of $z_{i,A}(p)$, according to the configuration LP specified in the previous section. Second, it takes polynomially-many samples and computes the empirical average, rather than an exact expected value. This allows us to compute $\hat{\psi}$ in poly-time. Finally, we write a convex quadratic minimization program whose solution is a fixed-point of $\hat{\psi}$. Because we can minimize convex quadratic functions in poly-time, we can then find a fixed point of $\hat{\psi}$.

More precisely we proceed as follows:

1. For $s \in \{1, \ldots, S\}$, with $S = \mathrm{poly}(|M|, |N|, 1/\varepsilon)$, draw independent sets of samples of the valuations $(v_i^{(s)})_{i\in N}$.
2. For each set of samples $(v_i^{(s)})_{i\in N}$ find an optimal fractional allocation $x^{(s)} = (x_{i,A}^{(s)})_{i\in N, A\subseteq M}$, i.e., one that solves

$$\text{(LP)} \qquad \begin{aligned} \max_{x\geq 0} \sum_{i\in N} \sum_{A\subseteq M} & x_{i,A} \cdot v_i^{(s)}(A) \\ \text{s.t.} \sum_{A\subseteq M} & x_{i,A} \leq 1, \text{ for all } i \in N, \\ \sum_{i\in N} \sum_{A:j\in A} & x_{i,A} \leq k_j, \text{ for all } j \in M. \end{aligned}$$

3. For each $s = 1, \ldots, S$ define the functions $\hat{\psi}^{(s)} : \mathbb{R}^M_{\geq 0} \to \mathbb{R}^M_{\geq 0}$ as

$$\hat{\psi}_j^{(s)}(p) = \frac{1}{k_j} \sum_{i\in N} \sum_{A:j\in A} x_{i,A}^{(s)} \cdot \left[v_i^{(s)}(A) - \sum_{j'\in A} p_{j'} \right]_+, \text{ for each } j \in M, \tag{8}$$

and denote their average as $\hat{\psi} := \frac{1}{S}\sum_{s=1}^{S} \hat{\psi}^{(s)}$.

4. Find a fixed point of $\hat{\psi}$, i.e., a vector $\hat{p}$ such that $\hat{\psi}(\hat{p}) = \hat{p}$.

As said before, $\hat{\psi}$ does not exactly approximate ψ, but another function $\tilde{\psi} := \mathbb{E}(\hat{\psi})$. Notice that if we define

$$\tilde{z}_{i,A}(p) = \mathbb{E}\left(x_{i,A} \cdot \left[v_i(A) - \sum_{j\in A} p_j \right]_+ \right),$$

then $\tilde{\psi}$ is analogous to ψ as defined in Eq. (5), but using $\tilde{z}$ instead of z, i.e.,

$$\tilde{\psi}_j(p) = \frac{1}{k_j} \sum_{i \in N} \sum_{A \subset M : j \in A} \tilde{z}_{i,A}(p).$$

To prove Theorem 5.1, we show that (i) given a set of valuation functions, we can efficiently compute an optimal solution of the linear program (LP) using demand queries; (ii) with polynomially many samples, the function $\hat{\psi}$ approximates $\tilde{\psi}$ sufficiently well; (iii) we can efficiently compute a fixed point of $\hat{\psi}$; and (iv) a fixed point of $\hat{\psi}$ (and thus an approximate fixed point of $\tilde{\psi}$) gives a $(d+1+\varepsilon)$-approximation of OPT.

Even though the linear program (LP) has exponentially many variables, its dual has only $|M|+|N|$ variables. It turns out the demand queries provide a separation oracle for it, and therefore, it can be solved using the Ellipsoid method in polynomial time. For more details we refer to [NRTV07, Chapter 11.5.2]. This completes step (i). For each of the other three steps we prove a separate lemma.

Lemma 5.2. *Using $S = \mathrm{poly}(|M|, |N|, 1/\varepsilon)$ samples we can guarantee that with probability $1-\varepsilon$ we have that $\sum_{j \in M} |\hat{\psi}_j(p) - \tilde{\psi}_j(p)| \le \varepsilon \cdot OPT/(|M| \cdot |N|)$, for all $p \in [0, v_{\max}]^M$.*

Proof. Consider the following discretization of $[0, v_{\max}]^M$. In each coordinate we take multiples of $\delta = \varepsilon \cdot OPT/(4 \cdot |M|^3 \cdot |N|)$, i.e., we consider vectors in $\mathcal{P} = \{i \cdot \delta : i \in \mathbb{N}\}^M \cap [0, v_{\max}]^M$. Recall that $\tilde{\psi} = \mathbb{E}(\hat{\psi})$. For any given $p \in \mathcal{P}$, $\lambda > 0$, $j \in M$, and number of samples S, an additive Chernoff bound indicates that

$$\mathbb{P}(|\hat{\psi}_j(p) - \mathbb{E}(\hat{\psi}_j(p))| > \lambda) \le 2 \exp\left(-\frac{2 \cdot S \cdot \lambda^2}{v_{\max}^2}\right).$$

Taking a union bound over all $j \in M$ and all $p \in \mathcal{P}$, we have that $\sum_{j \in M} |\hat{\psi}_j(p) - \mathbb{E}(\hat{\psi}_j(p))| \le |M| \cdot \lambda$ for all $p \in \mathcal{P}$ with probability at least

$$\begin{aligned}
&1 - |M| \cdot |\mathcal{P}| \cdot 2\exp\left(-\frac{2 \cdot S \cdot \lambda^2}{v_{\max}^2}\right) \\
&= 1 - |M| \cdot \left(\frac{v_{\max}}{\delta}\right)^{|M|} \cdot 2\exp\left(-\frac{2 \cdot S \cdot \lambda^2}{v_{\max}^2}\right) \\
&= 1 - \varepsilon \cdot \exp\left(\log\frac{|M|}{\varepsilon} + |M| \cdot \log\frac{v_{\max}}{\delta} - S \cdot 2 \cdot \frac{\lambda^2}{v_{\max}^2}\right) \qquad (9)
\end{aligned}$$

Now take any vector $p \in [0, v_{\max}]^M$. By the definition of $\mathcal{P}$, there is a vector $\hat{p} \in \mathcal{P}$ such that $||p - \hat{p}||_1 \le |M| \cdot \delta$. It is easy to check from the definition of $\hat{\psi}$ in Eq. (8) and the constraints that $x^{(s)}$ satisfies in LP that for all $j \in M$, $|\hat{\psi}_j(p) - \hat{\psi}_j(\hat{p})| \le ||p - \hat{p}||_1$. Therefore, $\hat{\psi}$ and $\mathbb{E}(\hat{\psi})$ are $|M|$-lipschitz functions. By the triangle inequality, we have that

$$\begin{aligned}
||\hat{\psi}(p) - \mathbb{E}(\hat{\psi}(p))||_1 &\le ||\hat{\psi}(p) - \hat{\psi}(\hat{p})||_1 + ||\hat{\psi}(\hat{p}) - \mathbb{E}(\hat{\psi}(\hat{p}))||_1 + ||\mathbb{E}(\hat{\psi}(\hat{p})) - \mathbb{E}(\hat{\psi}(p))||_1 \\
&\le 2 \cdot |M|^2 \cdot \delta + |M| \cdot \lambda. \qquad (10)
\end{aligned}$$

Now, taking $\lambda = \varepsilon \cdot OPT/(2 \cdot |M|^2 \cdot |N|)$ and replacing in Eq. (10), we obtain that $||\hat{\psi}(p) - \mathbb{E}(\hat{\psi}(p))||_1$ is at most $\varepsilon \cdot OPT/(|M| \cdot |N|)$ for all $p \in [0, v_{\max}]^M$; with probability at least the expression in Eq. (9). Assuming $v_{\max}/OPT \le \mathrm{poly}(|M|, |N|)$, we can make make the probability in Eq. (9) larger than $1 - \varepsilon$ by taking $S = \mathrm{poly}(|M|, |N|, |1/\varepsilon)$. □

Lemma 5.3. *If $S = \mathrm{poly}(|M|, |N|, 1/\varepsilon)$ we can compute a fixed point of $\hat{\psi}$ in time $\mathrm{poly}(|M|, |N|, 1/\varepsilon)$.*

Proof. Recall that p is a fixed point of $\hat{\psi}$ if for all $j \in M$,

$$p_j = \hat{\psi}_j(p) = \frac{1}{S}\sum_{s=1}^{S}\sum_{i\in N}\sum_{A:j\in A}\frac{1}{k_j}\cdot x_{i,A}^{(s)}\cdot\left[v_i^{(s)}(A) - \sum_{j'\in A} p_{j'}\right]_+ .$$

Note that in this sum, the only non-zero terms are those such that $x_{i,A}^{(s)} > 0$. A basic solution for the LP has at most $|M|+|N|$ non-zero variables, so there are at most $S\cdot(|M|+|N|) = \mathrm{poly}(|M|, |N|, 1/\varepsilon)$ combinations of indices such that $x_{i,A}^{(s)} > 0$. Denote by E the set of such indices, i.e., $E = \{(i, A, s) : i \in N, A \subseteq M, 1 \le s \le S, \text{ and } x_{i,A}^{(s)} > 0\}$.

Now, for a vector p, define

$$y_e := \sqrt{x_{i,A}^{(s)}}\cdot\left[v_i^{(s)}(A) - \sum_{j\in A} p_j\right]_+, \text{ for all } e = (i, A, s) \in E. \tag{11}$$

If p is a fixed point, then it satisfies

$$p_j = \sum_{\substack{(i',A',s')=e':\\ e'\in E, j\in A'}} \frac{\sqrt{x_{i',A'}^{(s')}}}{S\cdot k_j}\cdot y_{e'}. \tag{12}$$

By replacing p_j back in Eq. (11), we have that p is a fixed point if and only if $y = (y_e)_{e\in E}$ satisfies

$$y_e = \sqrt{x_{i,A}^{(s)}}\cdot\left[v_i^{(s)}(A) - \sum_{j\in A}\sum_{\substack{(i',A',s')=e':\\ e'\in E, j\in A'}} \frac{\sqrt{x_{i',A'}^{(s')}}}{S\cdot k_j}\cdot y_{e'}\right]_+, \quad \text{for all } e = (i, A, s) \in E. \tag{13}$$

We write a quadratic program with variables $(y_e)_{e\in E}$ whose optimal solutions correspond to solutions of Eq. (13).

$$\text{(QP)} \qquad \min_y \sum_{e=(i,A,s)\in E} y_e\cdot\left(y_e - \sqrt{x_{i,A}^{(s)}}\cdot\left(v_i^{(s)}(A) - \sum_{j\in A}\sum_{\substack{(i',A',s')=e':\\ e'\in E, j\in A'}} \frac{\sqrt{x_{i',A'}^{(s')}}}{S\cdot k_j}\cdot y_{e'}\right)\right)$$

s.t.

$$y_e \ge \sqrt{x_{i,A}^{(s)}}\cdot\left(v_i^{(s)}(A) - \sum_{j\in A}\sum_{\substack{(i',A',s')=e':\\ e'\in E, j\in A'}} \frac{\sqrt{x_{i',A'}^{(s')}}}{S\cdot k_j}\cdot y_{e'}\right), \; e = (i, A, s) \in E \tag{14}$$

$$y_e \ge 0, \; e \in E. \tag{15}$$

To see that it suffices to optimize this quadratic program, take first a vector y that satisfies Eq. (13) (note that such a vector must exist, by Brouwer's fixed-point theorem). It is immediately implied by Eq. (13) that y satisfies both Eqs. (14) and (15). Moreover, it is also evident that for all $e \in E$ one of the two constraints must be tight, implying that the objective function must take a value of 0. Notice that the objective function is necessarily non-negative for feasible solutions, so y is an optimal solution. Observe also that for any optimal solution y' to the quadratic program, because the objective function takes a value of zero it must be the case that for every $e \in E$, at least one of Eqs. (14) and (15) is tight. This directly shows that y' satisfies Eq. (13).

We finally argue that the quadratic program is convex and hence can be solved in polynomial time [KTK79]. To do so, it suffices to argue that the objective function can be written in the form $b^T y + y^T(B^T B + I)y$ for some vector b and matrix B. We define $B \in \mathbb{R}^{M \times E}$ by

$$B_{j,e=(i,A,s)} := \sqrt{\frac{x_{i,A}^{(s)}}{S \cdot k_j}} \cdot \mathbb{1}_{j \in A}.$$

Now, for $e = (i, A, s)$ and $e' = (i', A', s')$ observe that

$$(B^T B)_{e,e'} = \sum_{j \in M} \frac{\sqrt{x_{i,A}^{(s)} \cdot x_{i',A'}^{(s')}}}{S \cdot k_j} \cdot \mathbb{1}_{j \in A, j \in A'} = \sqrt{x_{i,A}^{(s)}} \cdot \left(\sum_{j \in A \cap A'} \frac{\sqrt{x_{i',A'}^{(s')}}}{S \cdot k_j} \right).$$

From this is is straightforward to see that all the nonlinear terms in the objective of QP can be written as $y^T(B^T B + I)y$ as we wanted to show. □

Lemma 5.4. *If $\hat{p} \in [0, v_{\max}]^M$ is such that $\sum_{j \in M} |\hat{p}_j - \tilde{\psi}_j(p)| \le \varepsilon \cdot OPT/(|M| \cdot |N|)$, then*

$$(d + 1 + O(\varepsilon)) \cdot ALG(\hat{p}) \ge OPT.$$

Proof. We simply re-do the proof of Theorem 3.1, but replacing with the approximate fixed point. Thus, we take $\hat{p} \in [0, v_{\max}]^M$ such that $\sum_{j \in M} |\hat{p}_j - \tilde{\psi}_j(p)| \le \varepsilon \cdot OPT/(|M| \cdot |N|)$, and replace in the bound of Lemma 4.1 We obtain that

$$\begin{aligned} ALG(\hat{p}) &\ge \min_{C \subseteq M} \left\{ \sum_{j \in C} k_j \cdot \hat{p}_j + \sum_{i \in N} \sum_{A \subseteq \bar{C}} \tilde{z}_{i,A}(\hat{p}) \right\} \\ &\ge \sum_{i \in N} \sum_{A \subseteq M} \tilde{z}_{i,A}(\hat{p}) - \frac{\varepsilon \cdot OPT}{|M| \cdot |N|} \cdot \max_{j \in M} k_j \\ &\ge \sum_{i \in N} \sum_{A \subseteq M} \tilde{z}_{i,A}(\hat{p}) - \frac{\varepsilon \cdot OPT}{|M|}. \end{aligned}$$

The last inequality comes from the fact that we can assume without loss of generality that $\max_{j \in M} k_j \le |N|$, since a buyer buys at most one copy of each item. Then, we use the upper bound of Lemma 4.2.

$$\begin{aligned} OPT &\le \sum_{j \in M} k_j \cdot \hat{p}_j + \sum_{i \in N} \sum_{A \subseteq M} \tilde{z}_{i,A}(\hat{p}) \\ &\le (d+1) \sum_{i \in N} \sum_{A \subseteq M} \tilde{z}_{i,A}(\hat{p}) + \frac{\varepsilon \cdot OPT}{|M| \cdot |N|} \cdot \max_{j \in M} k_j. \end{aligned}$$

This implies that

$$(1 - \varepsilon) \cdot OPT \le (d+1) \cdot \sum_{i \in N} \sum_{A \subseteq M} \tilde{z}_{i,A}(\hat{p}).$$

Putting together the lower bound on $ALG(\hat{p})$ and the upper bound on OPT we conclude that

$$(d + 1 + O(\varepsilon)) ALG(\hat{p}) \ge OPT.$$ □

The proof of the theorem is straightforward from the lemmas: we take $S = \text{poly}(|M|, |N|, \varepsilon)$ samples of the valuations, as required by Lemma 5.2 so that $\hat{\psi}$ is a good approximation of $\tilde{\psi}$. For each sample, we solve (LP) in polynomial time, so we can calculate $\hat{\psi}$ in polynomial time. Then, by Lemma 5.3, we can compute a fixed point of $\hat{\psi}$. Finally, taking the computed fixed point as prices, we get a $(d + 1 + O(\varepsilon))$-approximation of OPT, by Lemma 5.4.

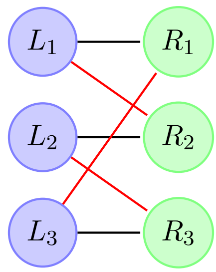

Figure 1: Example of a bipartite graph in which, when all edges have the same weight, no pricing scheme can guarantee obtaining more than 2/3 of the optimal solution.

6 Deterministic Single-minded Valuations

In this section, we consider the special case where there is a single copy of each item (i.e., $k_i = 1$ for all $i \in M$), buyers' valuations are deterministic, and buyers are *single-minded*. The latter means each buyer i has a set A_i, with $|A_i| \le d$, such that $A_i \not\subseteq B \Rightarrow v_i(B) = 0, A_i \subseteq B \Rightarrow v_i(B) = v_i(A_i)$. The problem of maximizing the welfare of an allocation in this context can be seen as the classic combinatorial problem of hypergraph matching with hyperedges of size at most d, where the buyers correspond to the hyperedges and the items are the vertices. Indeed, in an optimal allocation for this setting buyer i either gets A_i or $\emptyset$, implying that maximizing the welfare of the allocation is equivalent to finding a subset of pairwise disjoint A_i's of maximum total valuation. As this is a traditional problem, in the rest of this section we will refer to hypergraphs, hyperedges and vertices, rather than buyers and items, using the usual notation $G = (V, E)$ and denoting by $w(e)$ the valuation (or weight) of the hyperedge e.

6.1 Matching in graphs: $d = 2$

We first focus on the traditional matching problems, showing that using prices has limits even for this scenario. A similar discussion appears in [CAEFF16], though not for single-minded valuations. While our results are simple, and moreover Lemma 6.1 is implied by [CAEFF16]), we describe them here in full since we believe it helps complete the picture of the single-minded case. In particular, we show in Lemmas 6.1 and 6.3, there are instances in which no pricing scheme can guarantee recovering more than 2/3 of the optimal solution. This is true even if the graph is bipartite or if there is a unique optimal matching; on the other hand, if both conditions are fulfilled — the graph is bipartite and there is a unique optimal matching — using the dual prices leads precisely to such optimal solution.

Lemma 6.1. *Prices cannot guarantee obtaining more than* 2/3 *of the optimal matching, even if the graph is bipartite.*

Proof. Consider the graph depicted in Fig. 1, in which all edges have unit weight. There are two optimal solutions, given by the black and the red perfect matchings; no perfect matching can be constructed using arcs of different colors. Assume we have prices that are able to build an optimal solution (i.e., include three edges) regardless of the order in which the edges arrive. This implies that for at least one of the optimal solutions, all the edges will be included if their vertices are available when they arrive. Without loss assume this is the case for the black matching, i.e. for $i = 1, 2, 3, p_{L_i} + p_{R_i} \le 1$.

On the other hand, we need to prevent the red edges to be included if they appear; to see why this is necessary, consider for instance the case in which the edge (L_1, R_2) is not discarded when appearing first. Then, if the edge (L_3, R_3) appears second, no more edges could be added. To preclude this, we need to impose that for $i = 1, 2, 3$, $p_{L_i} + p_{R_{(i+1) \bmod 3}} > 1$. A contradiction follows by adding these and the previous three inequalities.

Finally, if, for instance, all vertex prices are $1/2$, exactly two edges will be added regardless of the order in which they appear. $\square$

In the case of bipartite graphs, it is natural to consider the usual linear programming formulation, since it has integer optimal solutions. The following lemma shows that when we require the additional hypothesis that there is a unique optimal matching. the prices given by the optimal solution of the dual problem lead to that optimal assignment.

Lemma 6.2. *If the graph $G = (V, E)$ is bipartite and has a unique optimal matching, then such a matching is obtained using the dual prices.*

Proof. Because the graph is bipartite, the problem reduces to solving the following linear program: $\max\{\sum_{e\in E} x_e w(e) : \sum_{e\in\delta(v)} x_e \leq 1 \text{ for all } v \in V, x \geq 0\}$; which has an integral optimal solution. Because there is only one optimal matching, the LP has a unique optimal solution $(x_e^*)_{e\in E}$. Consider the prices $(p_u^*)_{u\in V}$ corresponding to an optimal dual solution, satisfying strict complementary slackness.

Consider an edge $e = (u, v)$ that is not part of the optimal matching. Hence, the corresponding primal variable takes the value $x_e^* = 0$. By complementary slackness, the corresponding dual constraint is not tight, i.e. $p_u^* + p_v^* > w(e)$. This last condition implies that buyer e will not buy the edge upon arrival. On the other hand, if e is part of the optimal solution, the corresponding dual constraint must be tight (again due to strict complementary slackness), so that those buyers will choose to buy. $\square$

The assumption of a unique solution is crucial for the dual prices to be useful. Indeed, when there is more than one solution, using the dual prices can be arbitrarily inefficient. Indeed, consider the same example depicted in Figure 1, but modify the weight of the edges $f = (L_1, R_1)$ and $g = (L_2, R_3)$ to be ε, so that that the optimal solution has value $2 + \varepsilon$. On the other hand, consider an edge $e = (u, v)$ and the resulting dual prices p_u, p_v: complementary slackness now states that we have $p_u + p_v = w(e)$ iff e is part of any optimal solution. Edge f is part of the black optimal solution, and edge g is part of the red, hence those edges will be bought if the corresponding vertices are available when they appear. In particular, if they are the first two edges to appear, then they will both be in the final solution, and no other edge can be added, leading to a final weight of 2ε.

However, in general graphs, even the uniqueness assumption is not enough. Indeed we have the following result.

Lemma 6.3. *Prices cannot guarantee obtaining more than $2/3$ of the optimal matching in a general graph, even if there is only one optimal matching.*

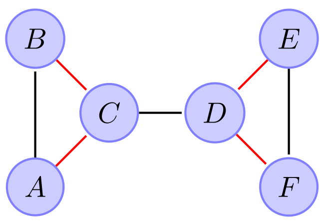

Figure 2: Example of a graph in which, when all edges have the same weight, there is a unique optimal matching but no pricing scheme can guarantee obtaining more than $2/3$ of its weight.

Proof. Consider the graph depicted in Fig. 2, where every edge has unit weight. The optimal matching is given by the three black edges with total value of 3. On the other hand, if any red edge enters the solution, the resulting total weight will be at most 2. We now show that any pricing scheme in which every black edge is willing to buy will also include at least one red edge if it comes first. Let $(p_i)_{i=A,\ldots,F}$ prices such that for every black edge, the sum of the involved vertices is lower than 1. In particular, we have that $p_C + p_D \leq 1$, so without loss of generality we assume that $p_C \leq 1/2$. If $p_B \leq 1/2$ as well, then the red edge (B, C) will want to buy and the proof is complete. Otherwise, if $p_B > 1/2$, it implies that $p_A \leq 1/2$ because the black edge (A, B) wants to buy. But this implies that the red edge (A, C) will buy if appearing first.

Finally, if all vertex prices are $1/2$, then it is straightforward to see that at least two edges will be added regardless of the order in which they appear. □

In general, there are item prices that guarantee obtaining at least half of the optimal welfare. This is achieved by splitting the weight of the edges of an optimal matching uniformly between the two corresponding vertices. We present this result in Lemma 6.6 for general d.

6.2 Hypergraph matching: $d > 2$

We begin this section proving two negative results. First we show an upper bound of $(1 + o(1)) \cdot \sqrt{\frac{1}{d}}$ on the fraction of the optimal solution that can be guaranteed with prices. We then show a specific bound for the case $d = 3$, in which we cannot guarantee obtaining more than $1/2$ of the optimal welfare. Finally, we provide a pricing scheme that always obtains at least $1/d$ of the optimal welfare.

Lemma 6.4. *Prices cannot guarantee welfare more than an $(1 + o(1)) \cdot \sqrt{\frac{1}{d}}$ fraction of the optimal welfare, even if the arrival order is known. More preicsely, for any q such that $q^2 - q + 1 \leq d$, prices cannot guarantee more than a $\frac{q}{q^2 - q + 1}$ approximation.*

Proof. Our example is based on constructions for finite projective planes; namely, we will use the fact that if $q - 1$ is a prime power there exists a hypergraph on $q^2 - q + 1$ vertices with $q^2 - q + 1$ hyperedges that are q-regular, q-uniform and intersecting, i.e. every pair of hyperedges has at least one shared vertex (see, e.g., [Hal98, Chapter 12] for a reference).

To build our example, we will assume that for each hyperedge there exists a corresponding buyer interested in exclusively that subset of items with a total valuation of q. We will also add one buyer whose only subset of interest is the entire set of items, with a valuation of $d = q^2 - q + 1$. Note that clearly the optimal welfare attainable is $q^2 - q + 1$.

It hence suffices to show that prices cannot achieve welfare greater than q. Assume the buyer interested in the entire set of items arrives last. Note that if there is any edge e such that the sum of the prices of the vertices in e is at most than q, we are guaranteed welfare at most q. However, if the sum of the prices of the vertices in every hyperedge is more than q, because our graph is q-uniform that means the sum of the prices of *all* vertices is more than $q^2 - q + 1$, meaning the final buyer would not select anything and the welfare attained is zero. Hence, the total welfare attainable by prices is at most a

$$\frac{q}{q^2 - q + 1} = (1 + o(1)) \cdot \frac{1}{\sqrt{q^2 - q + 1}}$$

fraction of the optimum.

Finally, if d cannot be written as $q^2 - q + 1$, we replicate the same construction for the largest $d' < d$ that can, and the result holds. □

When $d = 3$, taking $q = 2$ in the above proof demonstrates no prices achieve better than a 2/3-approximation. Below, we provide a tighter bound for this special case.

Lemma 6.5. *When $d = 3$, no prices exist obtaining a better than $1/2$-approximation.*

Proof. Consider a hypergraph $G = (V, E)$ with $V = \{1, 2, 3, 4, 5, 6\}$ and the following hyperedges with unit weight:

$$\{1, 2, 3\}, \{4, 5, 6\}, \{1, 2, 4\}, \{1, 3, 5\}, \{2, 5, 6\}, \{3, 4, 6\}.$$

First, note that there is a perfect matching (of weight 2), given by the two first hyperedges. Also note that each of the remaining hyperedges intersect all other hyperedges, thus only one of them could be included in a feasible solution. Therefore, it suffices to prove that there is no pricing scheme in which the first two edges ($\{1,2,3\},\{4,5,6\}$) want to buy but all the others do not.

Let us assume that there are prices $p_1,\dots,p_6$ that achieve the aforementioned property. For the first two edges to be taken when they appear, we need $p_1+p_2+p_3 \le 1$ and $p_4+p_5+p_6 \le 1$. To prevent the other edges to buy when they appear, we need the sum of the corresponding vertices to be strictly greater than 1, hence we obtain four additional inequalities. Adding up all the six inequalities, we obtain $\sum_{i=1}^{6} p_i > 2$. And the exact opposite result is obtained adding the first two inequalities. $\square$

We now provide our positive result. Consider a hypergraph $G=(V,E)$, with weights $(w(e))_{e\in E}$. To define the prices, take an optimal matching given by the hyperedges $OPT_1,\dots,OPT_\ell$. For each $a \in OPT_j$, define $p_a = w(OPT_j)/d$. The prices of the items not covered by the optimal solution are set to ∞. The following simple result shows that these prices obtain at least a fraction $1/d$ of the optimal welfare.

Lemma 6.6. *Consider prices defined as above, and hyperedges arriving in an arbitrary order. Denote Q the set of edges that are bought. Then*

$$\sum_{e\in Q} w(e) \ge \frac{1}{d}\sum_{j=1}^{\ell} w(OPT_j)$$

Proof. First note that for each $e \in Q$, it must hold that

$$w(e) \ge \sum_{i\in e} p_i \tag{16}$$

As otherwise the buyer associated to e would have decided not to buy. Therefore:

$$\sum_{e\in Q} w(e) \ge \sum_{e\in Q}\sum_{i\in e} p_i \tag{17}$$

On the other hand, for each OPT_j in the optimal solution, there must be at least one vertex, with its corresponding price $w(OPT_j)/d$ that is covered by the edges in Q. To see this, note that there are two possible cases: either $OPT_j \in Q$ and all its vertices are covered, or $OPT_j \notin Q$, meaning that when OPT_j arrived, at least one of its vertices was not available, i.e., it was covered by an edge previously bought. The result follows directly, noting that in the right side of Eq. (17), we are summing at least once $w(OPT_j)/d$ for each $j = 1\dots,\ell$. $\square$

7 Prophet Inequalities for Single-minded Random Valuations

In this section, we continue to study the setting with single-minded buyers, but allow them to have *random* valuations. This captures the well-studied *prophet inequality* setting. Indeed, when $d = 2$, our result recovers the 1/3-approximate thresholding prophet inequality of [GW19] for bipartite matching under edge-arrivals and extends this to matching in general graphs. (Vertices correspond to items, and each edge with a random weight corresponds to a buyer interested in two items with a random valuation). In this section, we show that in this setting there does not exist a prophet inequality guaranteeing better than a 3/7-approximation for bipartite matching, improving on the previous upper bound of 4/9 given by [GW19]. We stress that this upper bound holds against any online algorithm, not just thresholding algorithms. We also note that this bound extends the 3/7 upper bound given by [EFGT20] for matching in *general* graphs to bipartite graphs.[5]

[5]The constructions are independent and require different techniques. It is interesting that the current upper bounds for general vs. bipartite graphs stand at the same number.

Theorem 7.1. *No prophet inequality for bipartite matching in the edge-arrival setting is better than 3/7-competitive.*

The remainder of this section proves Theorem 7.1. We construct a family of hard instances $\{G_{n,k}\}$, indexed by integers $k \geq 2$ and $n \geq 2$. Below, we specify the construction of $G_{n,k}$ including an ordering on the edges and distribution $W(e)$ for the weight of each edge e. $G_{n,k}$ consists of:

- Nodes: $\{\ell_0, \ldots, \ell_{n-1}, r_0, \ldots, r_{n-1}\}$, for $2n$ total nodes.
- Edge: $a := (\ell_0, r_0)$. $W(a) = k^3$ with probability $1/k^3$, and 0 otherwise.
- Edges: $b_i := (\ell_0, r_i)$, $i \geq 1$. $W(b_i) = \frac{(k-1)^{2i-1}+k^{2i-1}}{(2k-1)k^{2i-2}}$ with probability $1/k$, and 0 otherwise.
- Edges: $c_i := (\ell_i, r_0)$, $i \geq 1$. $W(c_i) = \frac{k-1}{k} \cdot \frac{(k-1)^{2i-1}+k^{2i-1}}{(2k-1)k^{2i-2}}$ with probability $1/k$, and 0 otherwise.
- Edges arrive in the order $(c_{n-1}, b_{n-1}, c_{n-2}, b_{n-2}, \ldots, c_1, b_1, a)$.

This is depicted in Figure 3. To gain some intuition for this construction, observe that the realized weight of the final edge a has expectation 1. For all other edges, if this edge realizes with non-zero weight and an optimal online algorithm accepts, the remaining expected gain can be computed by solving a single-item Bayesian selection problem for the remaining disjoint edges. The weights $\{W(b_i), W(c_i)\}$ are carefully set so that at each point, if the realized weight is non-zero, the gain of an optimal online algorithm from accepting is exactly 1.

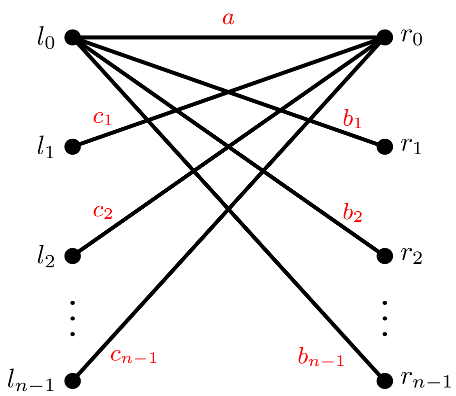

Figure 3: The structure of the hard instance

We first characterize the offline optimum. We let $OPT(W, G_{n,k})$ denote the expected performance of the optimum offline algorithm on the instance $(W, G_{n,k})$ and $ALG(W, G_{n,k})$ denote the expected performance of an online algorithm ALG.

Lemma 7.2. $\lim_{k\to\infty} \lim_{n\to\infty} OPT(W, G_{n,k}) = \frac{7}{3}$.

Proof. Observe first that maximal feasible sets can either take the single edge a, or two edges of the form $\{b_i, c_j\}$. Observe first that for all $i \geq 1$:

$$\frac{(k-1)^{2i-1}+k^{2i-1}}{(2k-1)k^{2i-2}} \leq \frac{k^{2i-1}+k^{2i-1}}{(2k-1)k^{2i-2}} = \frac{2k}{2k-1} \leq 2.$$

On the other hand, $k^3 \geq 8$. This means that whenever $W(a) > 0$, the offline optimum is a. Whenever $W(a) = 0$, the offline optimum is to accept the heaviest b edge together with the heaviest c edge. Observe next that for any $i \geq 1$, we have

$$\frac{(k-1)^{2i-1}+k^{2i-1}}{(2k-1)k^{2i-2}} > \frac{(k-1)^{2i+1}+k^{2i+1}}{(2k-1)k^{2i}}.$$

In particular, this means that b_i is the heaviest b edge if and only if $W(b_i) > 0$ *and* $W(b_j) = 0$ for all $j < i$. Hence, b_i is the largest b edge with probability $\left(1 - \frac{1}{k}\right)^{i-1} \cdot \frac{1}{k}$. From this we can directly conclude that the expected weight of the heaviest b edge is

$$\sum_{i=1}^{n-1} \left(\frac{k-1}{k}\right)^{i-1} \cdot \frac{1}{k} \cdot \frac{(k-1)^{2i-1} + k^{2i-1}}{(2k-1)k^{2i-2}} = \frac{1}{2k-1} \cdot \sum_{i=1}^{n-1} \frac{(k-1)^{3i-2} + k^{2i-1} \cdot (k-1)^{i-1}}{k^{3i-2}}.$$

Identical calculations (after multiplying all weights by $\frac{k-1}{k}$) yield that the expected weight of the heaviest c edge is

$$\frac{k-1}{k} \cdot \frac{1}{2k-1} \cdot \sum_{i=1}^{n-1} \frac{(k-1)^{3i-2} + k^{2i-1} \cdot (k-1)^{i-1}}{k^{3i-2}}.$$

Putting everything together, we conclude that

$$OPT(W, G_{n,k}) = \left(\frac{1}{k^3}\right) \cdot k^3 + \left(1 - \frac{1}{k^3}\right) \cdot \left(\frac{1}{k} \sum_{i=1}^{n-1} \frac{(k-1)^{3i-2} + k^{2i-1} \cdot (k-1)^{i-1}}{k^{3i-2}}\right).$$

We now can perform some straightforward manipulations; in particular, first observe that

$$\begin{aligned}
\sum_{i=1}^{\infty} \frac{(k-1)^{3i-2} + k^{2i-1} \cdot (k-1)^{i-1}}{k^{3i-2}} &= \sum_{i=1}^{\infty} \left(\frac{k-1}{k}\right)^{3i-2} + \sum_{i=1}^{\infty} \left(\frac{k-1}{k}\right)^{i-1} \\
&= \frac{k-1}{k \cdot (1 - (1 - 1/k)^3)} + k \\
&= \frac{k^3 - k^2}{k^3 - (k-1)^3} + \frac{k^4 - k(k-1)^3}{k^3 - (k-1)^3} \\
&= \frac{4k^3 - 4k^2 + k}{3k^2 - 3k + 1}.
\end{aligned}$$

Hence, for fixed k:

$$\lim_{n\to\infty} OPT(W, G_{n,k}) = \left(\frac{1}{k^3}\right) \cdot k^3 + \left(1 - \frac{1}{k^3}\right) \cdot \left(\frac{1}{k} \cdot \frac{4k^3 - 4k^2 + k}{3k^2 - 3k + 1}\right) = \frac{7k^5 + O(k^4)}{3k^5 + O(k^4)}.$$

From this we directly conclude

$$\lim_{k\to\infty} \lim_{n\to\infty} OPT(W, G_{n,k}) = 7/3$$

as claimed. □

Note that the original sum can be split into two because it is convergent. This can be seen through the two resulting sums in the first line, which for any fixed k are a geometrical series of a number lower than 1, namely $(k-1)/k$. Our next lemma considers the power of online algorithms for $G_{n,k}$ instances. In fact, we will see that for any $G_{n,k}$ instance an online algorithm cannot get a score better than 1 in expectation, which will lead to our lower bound of 7/3 on the competitive ratio.

Lemma 7.3. *For all online algorithms ALG and all $n, k \geq 2$, $ALG(W, G_{n,k}) \leq 1$.*

Proof. Fix some $n, k \geq 2$. We will prove the lemma with a sequence of intermediate claims. We first bound the expected reward from a set of edges $\{b_i, b_{i-1}, \dots, b_1\}$ revealed in that order. This is useful because once an algorithm has accepted some c edge, they cannot accept another c edge, or a, so they are left exactly with this optimization.

Claim 7.4. *The maximum expected reward that any online algorithm can achieve from the edges* $\{b_i, \ldots, b_1\}$ *(revealed in that order) is*

$$1 - \frac{k-1}{k} \cdot \frac{(k-1)^{2i-1} + k^{2i-1}}{(2k-1)k^{2i-2}} = \frac{k^{2i} - (k-1)^{2i}}{(2k-1)k^{2i-1}}.$$

Proof. We proceed by induction on i. Our base case is $i = 1$; in this case ALG clearly gets expectation at most $\mathbb{E}[W(b_1)] = \frac{1}{k}$ which indeed is precisely equal to the expression claimed.

Hence assume the result for some fixed $i \geq 1$; we will prove it for $i+1$. When ALG considers whether to accept or reject b_{i+1}, note that if it rejects it is guaranteed expectation at most $\frac{k^{2i}-(k-1)^{2i}}{(2k-1)k^{2i-1}}$ by the inductive hypothesis. Note that $W(b_{i+1})$ could take either a value of 0 or $\frac{(k-1)^{2i+1}+k^{2i+1}}{(2k-1)k^{2i}} > \frac{k^{2i}-(k-1)^{2i}}{(2k-1)k^{2i-1}}$. This implies that if $W(b_{i+1}) > 0$, ALG maximizes its expected reward from edges in $\{b_{i+1}, \ldots, b_1\}$ by accepting b_{i+1}. Furthermore, it is clear that if $W(b_{i+1})$ takes value 0, then it is optimal for ALG to reject it. This lets us conclude that the expectation ALG can attain is at most

$$\begin{aligned}
&\frac{1}{k} \cdot \frac{(k-1)^{2i+1} + k^{2i+1}}{(2k-1)k^{2i}} + \left(1 - \frac{1}{k}\right) \cdot \left(\frac{k^{2i} - (k-1)^{2i}}{(2k-1)k^{2i-1}}\right) \\
&\qquad = \frac{(k-1)^{2i+1} + k^{2i+1}}{(2k-1)k^{2i+1}} + \frac{k(k-1) \cdot (k^{2i} - (k-1)^{2i})}{(2k-1)k^{2i+1}} \\
&\qquad = \frac{k^{2i+2} - (k-1)^{2i+2}}{(2k-1)k^{2i+1}}
\end{aligned}$$

hence proving the inductive hypothesis. ∎

Note that identical calculations, after multiplying all weights by $\frac{k-1}{k}$, give the following as well:

Claim 7.5. *The maximum expected reward that any online algorithm can achieve from the edges* $\{c_i, \ldots, c_1\}$ *(revealed in that order) is*

$$\frac{k-1}{k} \cdot \left(1 - \frac{k-1}{k} \cdot \frac{(k-1)^{2i-1} + k^{2i-1}}{(2k-1)k^{2i-2}}\right) = \frac{k-1}{k} \cdot \frac{k^{2i} - (k-1)^{2i}}{(2k-1)k^{2i-1}}.$$

With these claims in place, we directly get an upper bound on the algorithm's performance based on the first edge it accepts in $\{c_1, c_2, \ldots, c_{n-1}\}$ or $\{b_1, b_2, \ldots, b_{n-1}\}$.

Claim 7.6. *If ALG first accepts some edge in* $\{c_1, c_2, \ldots, c_{n-1}\}$*, then its expected reward, conditioned on this event, is at most one.*

Proof. Let c_i be the first edge accepted by ALG. Note that the weight ALG gets from c_i is at most

$$\frac{k-1}{k} \cdot \frac{(k-1)^{2i-1} + k^{2i-1}}{(2k-1)k^{2i-2}}.$$

Note also that ALG cannot accept a or any other c_j (as this would not be a matching). Hence ALG can only accept at most one edge from $\{b_i, b_{i-1}, \ldots, b_1\}$. By Claim 7.4, the expected reward of the algorithm on the remaining edges is at most:

$$1 - \frac{k-1}{k} \cdot \frac{(k-1)^{2i-1} + k^{2i-1}}{(2k-1)k^{2i-2}}.$$

These two terms sum to one, yielding the result. ∎

We similarly have a symmetric claim for $\{b_1, b_2, \ldots, b_{n-1}\}$.

Claim 7.7. *If ALG first accepts some edge in* $\{b_1, b_2, \ldots, b_{n-1}\}$*, then its expected reward, conditioned on this event, is at most one.*

Proof. Let b_i be the first edge accepted by ALG. Note that the weight ALG gets from b_i is at most

$$\frac{(k-1)^{2i-1} + k^{2i-1}}{(2k-1)k^{2i-2}}.$$

Note also that ALG cannot accept a or any other b_j. Hence ALG can only accept at most one edge from $\{c_{i-1}, \ldots, c_1\}$. By Claim 7.5 the expected reward of the algorithm on the remaining edges is at most:

$$\frac{k-1}{k} \cdot \frac{k^{2i-2} - (k-1)^{2i-2}}{(2k-1)k^{2i-3}}.$$

Hence, in total ALG gets in expectation at most

$$\frac{(k-1)^{2i-1} + k^{2i-1}}{(2k-1)k^{2i-2}} + \frac{k-1}{k} \cdot \frac{k^{2i-2} - (k-1)^{2i-2}}{(2k-1)k^{2i-3}} = 1. \qquad \square$$

The only other case to consider is when ALG does not accept any edge in $\{b_1, b_2, \ldots, b_{n-1}\} \cup \{c_1, c_2, \ldots, c_{n-1}\}$. But conditioned on this, the algorithm can only accept a. As $\mathbb{E}[W(a)] = 1$, we also see that ALG gets expected value at most one, conditioned on this. To summarize, we have now shown that no matter what edge ALG accepts first from $\{c_{n-1}, b_{n-1}, \ldots, c_1, b_1\}$, it gets expected value at most one. We have also shown that if ALG accepts nothing from this set that it also gets expected value at most one. This completes the proof. $\square$

Summarizing, we know from Lemma 7.2 and Lemma 7.3 that for any fixed $\varepsilon > 0$, there exist some large n, k, such that $OPT(W, G_{n,k}) \geq 7/3 - \varepsilon$, while $ALG(W, G_{n,k}) \leq 1$ for all online algorithms. This implies an upper bound of $3/7$ on the attainable competitive ratio of online algorithms for bipartite matching, completing the proof of Theorem 7.1.

8 Conclusion and Future Directions

In this paper, we provided an efficiently computable $1/(d+1)$-approximate pricing algorithm for maximizing social welfare when buyers have arbitrary and random monotone valuations on subsets of at most d items. Although this approximation factor is tight in the worst case, numerous interesting directions for future work remain. In the special case where buyers are single-minded and have deterministic valuations, for $d = 2$ we have bounded the best attainable ratio of pricing algorithms in $[1/2, 2/3]$, so the exact value is yet to be found. It would furthermore be relevant to understand the asympotics of the optimal ratio for general d, which our results place in $[\sim 1/\sqrt{d}, 1/d]$. When buyers are single-minded but could have random valuations, our problem is closely related to the design of thresholding prophet inequalities. Many open problems remain in this area; we gave an upper bound for prophet inequalities for bipartite matching and note that there remain large gaps in known bounds for the optimal competitive ratio for matching in bipartite (and general) graphs.

Acknowledgements. We thank the two anonymous reviewers for carefully reading the paper and providing many constructive comments that greatly improved the presentation.

References

[AKS21] Sepehr Assadi, Thomas Kesselheim, and Sahil Singla. Improved truthful mechanisms for subadditive combinatorial auctions: Breaking the logarithmic barrier. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 653–661. SIAM, 2021.

[AS19] Sepehr Assadi and Sahil Singla. Improved truthful mechanisms for combinatorial auctions with submodular bidders. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 233–248. IEEE, 2019.

[BK19] Elizabeth Baldwin and Paul Klemperer. Understanding preferences:"demand types", and the existence of equilibrium with indivisibilities. *Econometrica*, 87(3):867–932, 2019.

[BLNPL14] Moshe Babaioff, Brendan Lucier, Noam Nisan, and Renato Paes Leme. On the efficiency of the walrasian mechanism. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 783–800, 2014.

[CAEFF16] Vincent Cohen-Addad, Alon Eden, Michal Feldman, and Amos Fiat. The invisible hand of dynamic market pricing. In *2016 ACM Conference on Economics and Computation (EC)*, pages 383–400. ACM, 2016.

[CC23] José R. Correa and Andrés Cristi. A constant factor prophet inequality for online combinatorial auctions. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 686–697. ACM, 2023.

[CHMS10] Shuchi Chawla, Jason D Hartline, David L Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 311–320, 2010.

[CM16] Shuchi Chawla and J. Benjamin Miller. Mechanism design for subadditive agents via an ex ante relaxation. In Vincent Conitzer, Dirk Bergemann, and Yiling Chen, editors, *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, Maastricht, The Netherlands, July 24-28, 2016*, pages 579–596. ACM, 2016.

[CZ17] Yang Cai and Mingfei Zhao. Simple mechanisms for subadditive buyers via duality. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 170–183. ACM, 2017.

[DFKL20] Paul Dutting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing nonstochastic inputs. *SIAM Journal on Computing*, 49(3):540–582, 2020.

[DKL20] Paul Dütting, Thomas Kesselheim, and Brendan Lucier. An o (log log m) prophet inequality for subadditive combinatorial auctions. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 306–317. IEEE, 2020.

[DNS05] Shahar Dobzinski, Noam Nisan, and Michael Schapira. Approximation algorithms for combinatorial auctions with complement-free bidders. In Harold N. Gabow and Ronald Fagin, editors, *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 610–618. ACM, 2005.

[Dob16] Shahar Dobzinski. Breaking the logarithmic barrier for truthful combinatorial auctions with submodular bidders. In Daniel Wichs and Yishay Mansour, editors, *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 940–948. ACM, 2016.

[EFGT20] Tomer Ezra, Michal Feldman, Nick Gravin, and Zhihao Gavin Tang. Online stochastic max-weight matching: prophet inequality for vertex and edge arrival models. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 769–787, 2020.

[FGL14] Michal Feldman, Nick Gravin, and Brendan Lucier. Combinatorial auctions via posted prices. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*, pages 123–135. SIAM, 2014.

[FGL16] Michal Feldman, Nick Gravin, and Brendan Lucier. Combinatorial walrasian equilibrium. *SIAM Journal on Computing*, 45(1):29–48, 2016.

[FSZ16] Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 1014–1033. SIAM, 2016.

[GW19] Nikolai Gravin and Hongao Wang. Prophet inequality for bipartite matching: merits of being simple and non adaptive. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 93–109, 2019.

[Hal98] Marshall Hall. *Combinatorial theory*, volume 71. John Wiley & Sons, 1998.

[KTK79] Mikhail K Kozlov, Sergei Pavlovich Tarasov, and Leonid Genrikhovich Khachiyan. Polynomial solvability of convex quadratic programming. In *Doklady Akademii Nauk*, volume 248, pages 1049–1051. Russian Academy of Sciences, 1979.

[KV12] Piotr Krysta and Berthold Vöcking. Online mechanism design (randomized rounding on the fly). In *International Colloquium on Automata, Languages, and Programming*, pages 636–647. Springer, 2012.

[KW12] Robert Kleinberg and Seth Matthew Weinberg. Matroid prophet inequalities. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 123–136, 2012.

[LW20] Renato Paes Leme and Sam Chiu-wai Wong. Computing walrasian equilibria: Fast algorithms and structural properties. *Mathematical Programming*, 179(1):343–384, 2020.

[NRTV07] Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[SC84] Ester Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability*, pages 1213–1216, 1984.

[SVW23] Raghuvansh R. Saxena, Santhoshini Velusamy, and S. Matthew Weinberg. An improved lower bound for matroid intersection prophet inequalities. In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference, ITCS 2023, January 10-13, 2023, MIT, Cambridge, Massachusetts, USA*, volume 251 of *LIPIcs*, pages 95:1–95:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023.

[Tre01] Luca Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 453–461, 2001.