Power of Posted-price Mechanisms for Prophet Inequalities

Kiarash Banihashem [*] MohammadTaghi Hajiaghayi [†] Dariusz R. Kowalski [‡]
Piotr Krysta [§] Jan Olkowski [¶]

Abstract

We study the power of posted pricing mechanisms for Bayesian online optimization problems subject to combinatorial feasibility constraints. When the objective is to maximize social welfare, the problem is widely studied in the literature on prophet inequalities. While most (though not all) existing algorithms for prophet inequalities are implemented using a pricing mechanism, whether or not this can be done in general is unknown, and was formally left as an open question by Dütting, Feldman, Kesselheim, and Lucier (FOCS 2017, SICOMP 2020). Understanding the power and limitations of posted prices is important from a mechanism design perspective because any posted price mechanism is truthful, and is also interesting in its own right as it can guide future research on prophet inequalities.

We show that any prophet inequality has an implementation using a posted price mechanism, thereby resolving the open question of Dütting et al. Given an algorithm for Bayesian online optimization, we show that it can be transformed, in a black-box manner, to a posted price algorithm that has the same or higher expected social welfare and preserves the distribution over the assigned outcomes. We further show how to implement our reduction efficiently under standard assumptions using access to a sampling oracle. As an immediate consequence, we obtain improved pricing-based prophet inequalities for maximum weight matching, resolving an open problem of Ezra, Feldman, Gravin and Tang (EC 2020, MOR 2022). Correa and Cristi (STOC 2023) proved recently an existence of prophet inequality with constant approximation ratio for online social welfare maximizing combinatorial auctions with subadditive valuations. They left as an open problem to provide a posted pricing based implementation of their algorithm. Our technique resolves this question in affirmative as well.

1 Introduction

In the past two decades, there has been a surge of interest in computer science community for studying the *prophet inequality* problem. In the simplest problem's version, an online algorithm tries to choose a large value from a sequence of n values $v_1, \dots, v_n$, where each v_i is drawn from a known distribution D_i over positive real numbers. The values are revealed sequentially, and upon observing each v_i, the algorithm makes an irrevocable decision of whether to accept it or not. The objective of the algorithm is to maximize the expected value of the selected value. The standard benchmark for the algorithm is the expected offline optimum, i.e., $\mathbb{E}[\max_i v_i]$. In this paper, we consider the generalized version of this problem with combinatorial feasibility constraints, like those of combinatorial auctions. Here, a set of n agents arrive in an online manner, and each agent possesses a valuation function v_i over an outcome space X_i. As before, the valuation functions are sampled independently from known distributions. The goal of the problem is to assign an outcome to each agent in a way that maximizes the expected social welfare, subject to some combinatorial constraint over the assigned outcomes.

As a simple model for online selection, the prophet inequality problem is a central problem in optimal stopping theory and algorithmic game theory with connections to posted-price mechanisms [30, 9, 44, 14], stochastic probing [29, 1], and delegation [34, 4]. As such, many works have studied prophet inequalities and their variations for different combinatorial constraints such as matroids, matchings, knapsacks [35, 2, 26, 39, 27, 24], and combinatorial auctions [25, 19, 20, 16]. Recently, several works have also considered the random arrival order model, referred to as *prophet secretary* [22, 21, 15, 12].

A common approach for designing algorithms for prophet inequalities is via posted-price mechanisms [30, 9]. In this approach, the algorithm sets a price $\pi(x)$ for each outcome before observing the realized valuation function v_i, and upon observing v_i, assigns the outcome with highest utility to the agent, i.e., $\arg\max_{x\in X_i}(v_i(x) - \pi(x))$. For the classical version of the problem one possible choice is to set the threshold of $\tau = \mathbb{E}[\max v_i]/2$ and accept the first element with $v_i \geq \tau$. The pricing based approach has been used to design algorithms for prophet inequalities with combinatorial constraints such as matroids, matroid intersection and knapsacks [35, 27], combinatorial auctions with submodular and XOS valuation functions [25, 19], as well as closely related settings such as the prophet secretary problem [22, 21, 15, 12].

[*] University of Maryland, College Park, Maryland, USA. `kiarash@umd.edu`.

[†] University of Maryland, College Park, Maryland, USA. Email: `hajiagha@umd.edu`.

[‡] Augusta University, Augusta, Georgia, USA. Email: `dkowalski@augusta.edu`.

[§] Augusta University, Augusta, Georgia, USA. Email: `pkrysta@augusta.edu`. Piotr Krysta is also affiliated with Computer Science Department, University of Liverpool, U.K.

[¶] University of Maryland, College Park, Maryland, USA. Email: `olkowski@umd.edu`.

In addition to their simplicity and intuitive nature, posted price mechanisms offer the advantage of being dominant-strategy incentive-compatible (DSIC). This means that each agent has a weakly-dominant strategy of revealing their valuation function truthfully. Given the significance of this property, it is natural to explore whether the pricing approach can be applied universally to prophet inequalities, meaning if it can be used for all instances of the problem. This question was formally raised as an open problem by Dütting, Feldman, Kesselheim, and Lucier (FOCS 2017, SICOMP 2020) [18, 19]

...A related question is whether there exist prophet inequalities that cannot be implemented using posted prices. Interestingly, we are not aware of any separation between the two so far. ...

In this work, we resolve this open problem by showing that all prophet inequalities can be implemented using posted prices. Specifically, we provide a black-box reduction that, given an input algorithm for the prophet inequalities problem, transforms it into a pricing based algorithm that preserves the distribution of the original algorithm over the assigned outcomes, and has greater or equal social welfare. Black-box approaches are popular in algorithmic game theory as they apply to a wide range of problems, and are generally not algorithm specific [33, 32, 10, 7, 17, 14]. We further show how to implement our posted price mechanism efficiently under standard assumptions via sampling access to the distributions of the valuation functions.

As a direct consequence of our findings, we achieve an improved pricing-based algorithm for matching prophets. Notably, in a recent work Ezra, Feldman, Gravin, and Tang (EC 2020, MOR 2022) [23, 24] obtained enhanced competitive ratios for matching prophets in both the vertex arrival and edge arrival settings using a non-pricing approach. They then posed the question of whether similar results could be achieved using a pricing approach (see the end of the abstract and Section 5 of their paper for discussion of the problem). Our result resolves this open problem affirmatively. In a recent breakthrough, Correa and Cristi (STOC 2023) [16] proved an existence of prophet inequality for online social welfare maximizing combinatorial auctions with subadditive valuations, with the first known constant approximation ratio. They left as an open problem to provide a posted pricing-based implementation of their algorithm. Our technique lets us resolve this question in affirmative as well.

1.1 Our results.

Black-box reduction. Our main result takes the form of a black-box reduction that transforms any algorithm for prophet inequalities into a pricing based algorithm with comparable guarantees. We state our result here and refer to Section 4 for a more formal version and the proof.

Theorem 1. *Let $\mathcal{A}^{\text{inp}}$ be an algorithm for the prophet inequalities problem, and let $\mathbb{E}\left[\, v(\mathcal{A}^{\text{inp}})\,\right]$ denote its expected social welfare. There exists a pricing based algorithm $\mathcal{A}^{\text{out}}$ with expected social welfare $\mathbb{E}\left[\,\mathbf{v}(\mathcal{A}^{\text{out}})\,\right] \geq \mathbb{E}\left[\,\mathbf{v}(\mathcal{A}^{\text{inp}})\,\right]$. Furthermore, $\mathcal{A}^{\text{out}}$ and $\mathcal{A}^{\text{inp}}$ have the same distribution over the assigned outcomes. Formally, letting $\mathcal{A}_i \in X_i$ denote the assignment made by the algorithm $\mathcal{A}$ to agent i,*

$$\Pr\left[\,(\mathcal{A}_1^{\text{inp}}, \ldots, \mathcal{A}_n^{\text{inp}}) = (x_1, \ldots, x_n)\,\right] = \Pr\left[\,(\mathcal{A}_1^{\text{out}}, \ldots, \mathcal{A}_n^{\text{out}}) = (x_1, \ldots, x_n)\,\right] \tag{1.1}$$

for all $(x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n$.

We note that Equation (1.1) immediately implies that the reduction preserves any constraint on the distribution of the assigned outcomes that is satisfied by $\mathcal{A}^{\text{inp}}$. This includes combinatorial constraints as a special case; assuming $(\mathcal{A}_1^{\text{inp}}, \ldots, \mathcal{A}_n^{\text{inp}}) \in \mathcal{F}$ with probability 1 for some $\mathcal{F} \subseteq X_1 \times \cdots \times X_n$, the condition ensures that $(\mathcal{A}_1^{\text{out}}, \ldots, \mathcal{A}_n^{\text{out}}) \in \mathcal{F}$ with probability 1 as well. The condition is more general however and it also implies the preservation of probabilistic constraints over the assigned outcomes. For instance, specializing to combinatorial auctions, we can preserve constraints of the type "the first agent is allocated an item with probability $\geq 1/2$". We also note that while the optimal policy in combination with VCG payments (e.g., see Correa and Cristi [16]) also leads to a pricing based algorithm, the approach would change the distribution over outcomes and is inefficient.

In order to prove the theorem, for each arriving agent i, we start by considering the assignment that the input algorithm would have made for the agent based on the already assigned outcomes. This takes the form of a (randomized) mapping $\overline{A}_i^{\text{inp}}$ from the space of valuations functions V_i to the set of outcomes X_i. We refer to these mappings as *assignment rules*. Given $\overline{A}_i^{\text{inp}}$, we propose an alternative assignment rule A_i^{out} that preserves the assignment probability of each outcome, i.e.,

$$\Pr_{v_i \sim D_i}\left[A_i^{\text{out}}(v_i) = x\right] = \Pr_{v_i \sim D_i}\left[\overline{A}_i^{\text{inp}}(v_i) = x\right] \text{ for all } x \in X_i \;. \tag{1.2}$$

We show that as long as the new assignment rules satisfy the constraint (1.2), the algorithm satisfies the constraint (1.1). Our proof will crucially rely on the fact that $\overline{A}_i^{\text{inp}}$ is calculated based only on the assigned outcomes $A_1^{\text{inp}}, \ldots, A_{i-1}^{\text{inp}}$, and not on the observed valuation functions $v_1, \ldots, v_{i-1}$. We further show that the "optimal" choice of an assignment rule, i.e., the one maximizing $\mathbb{E}_{v_i \sim D_i}\left[A_i^{\text{out}}(v_i)\right]$ subject to (1.2), naturally leads to a pricing based mechanism. The crux of our analysis is a duality based argument: We formulate the problem of finding the optimal assignment rule as an optimal transport linear program (with possibly an infinite number of constraints), and show that strong duality for this linear program implies a pricing-based solution. We use this result, together with constraint (1.1), to show that $\mathbb{E}\left[v(\mathcal{A}^{\text{out}})\right] \geq \mathbb{E}\left[v(\mathcal{A}^{\text{inp}})\right]$. We refer to Section 4 for more details. As we will see, the assignment rule $\overline{A}_i^{\text{inp}}$ does not need to be calculated explicitly by the algorithm, and we will only require the probabilities $\Pr_{v_i \sim D_i}\left[\overline{A}_i^{\text{inp}}(v_i) = x\right]$. The theorem resolves the open problem of Dütting et al. [18, 19].

Computational efficiency. While the above results are about the existence of an algorithm, they do not say anything about efficiency. The reason for this is that the approach is based on a linear program where the number of constraints can be as large as the support of the agents' distribution over valuation functions, i.e., D_i. As such, while the runtime of the algorithms is polynomial in the length of the input, this can be impractical for continuous distributions, where the input size is not finite. To deal with this issue, we show how the above strategy can be implemented efficiently under standard assumptions using access to a sampling oracle from the distributions D_i. We state our result here and refer to Section 6 for a more formal version and the proof.

Theorem 2. *Let $\mathcal{A}^{\text{inp}}$ be an algorithm for the prophet inequalities problem with expected social welfare $\mathbb{E}\left[v(\mathcal{A}^{\text{inp}})\right]$ as in Theorem 1, let $\mathcal{F} \subseteq X_1 \times \cdots \times X_n$ be a combinatorial constraint satisfied by $\mathcal{A}^{\text{inp}}$, and let $\epsilon > 0$ be an arbitrary parameter. Given sampling access to the distributions D_i, under standard assumptions, there exists a pricing based algorithm $\mathcal{A}^{\text{out}}$ such that $\mathbb{E}\left[v(\mathcal{A}^{\text{out}})\right] \geq (1-\epsilon) \cdot \mathbb{E}\left[v(\mathcal{A}^{\text{inp}})\right]$. Furthermore, the joint distribution of the assignments for $\mathcal{A}^{\text{inp}}$ and $\mathcal{A}^{\text{out}}$ is the same up to a factor of $1-\epsilon$, and the algorithm uses $\text{poly}(n, \max_i |X_i|, \frac{1}{\varepsilon})$ samples and runs in $\text{poly}(n, \max_i |X_i|, \frac{1}{\varepsilon})$ time. Additionally, the algorithm is feasible under the combinatorial constraint $\mathcal{F}$.*

We note that algorithms based on sampling have previously been explored for prophet inequalities [3, 13, 40, 8]. The setting we consider is fairly general however and we make minimal assumptions on the input algorithm, while existing approaches are in more restricted settings and are generally algorithm specific. As such, our approach requires new techniques which may be of independent interest. We here provide an overview of our proof and refer to Section 6 for more details.

For each agent i, we effectively estimate the distribution D_i using a sufficient (but polynomial) number of samples, obtaining an empirical distribution $\widehat{D}_i$. We then show that solving the optimization problem in Theorem 2 on the estimated distribution leads to an almost optimal solution to the original problem. The main idea behind the proof is a uniform convergence argument based on bounding the VC-dimension of all pricing based algorithms. The standard way to do this is to uniformly bound, for all pricing based algorithms, the difference in the objective of the true optimization problem and the estimated optimization problem. This is difficult to do however because the objective depends on the input algorithm in a complicated way. Additionally, it is not clear how the difference in objective affects the probability distribution over the assigned outcomes, i.e., Equation (1.2), which is a central part of our proof of Theorem 1. Crucially, Equation (1.2) is necessary for the proof that the social welfare does not decrease, i.e., $\mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{out}})\right] \geq \mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{inp}})\right]$.

To deal with this issue, we take an indirect approach that effectively works in the dual space. We first uniformly bound, for all pricing assignment rules A and outcomes $x \in X_i$, the difference between $\Pr_{v \sim D}[A(v) = x]$ for $D = D_i$ (i.e., sampling from the true distribution), and $D = \widehat{D}_i$ (i.e., sampling from the estimated distribution). We do this by viewing each assignment probability as the success rate of a classifier formed by an intersection of half-spaces, and bounding the VC-dimension of these half-spaces. This leads to an additive bound on the error in the approximation. An additive error is not well suited for our purposes however as it means that low probability events may be ignored (as their probability is rounded down to 0), even though they may constitute a large fraction of an algorithm's expected welfare. To deal with this obstacle, we further "soften" our probability estimates by mixing them with the uniform distribution with low probability, changing the error bounds into multiplicative bounds. We use the multiplicative bounds to show that the algorithm approximately preserves both the distribution over outcomes and the expected social welfare. We refer to Section 6 for more details.

Applications. We can apply our main result to special cases of prophet inequalities found in the literature. Specifically, we obtain improved pricing based algorithms prophet inequalities with a matching constraint (for both edge arrival and vertex arrival models), resolving the open problem of Ezra et al. [23]. We also apply our reduction to online social welfare maximizing combinatorial auctions with subadditive valuations. In a recent breakthrough, Correa and Cristi [16] proved an existence of prophet inequality for this problem with the first known constant approximation ratio. They left as an open problem to provide a posted pricing based implementation of their algorithm. Our technique lets us resolve this question in affirmative as well. These applications are presented in Section 5.

Map of the paper. The remainder of the paper is organized as follows. In Section 2, we discuss the related work. In Section 3, we discuss the preliminaries of our model and establish the notation. In Section 4, we discuss our pricing algorithm, and prove our main result (Theorem 1). In Section 5, we discuss the applications of our result to prophet inequalities with matching constraints and combinatorial auctions. Finally, in Section 6, we show how our pricing algorithm can be efficiently implemented using sampling access to the distributions.

Remark 3. *While our focus here is on the prophet inequalities problem, our results immediately extend to the prophet secretary problem as each instance of the prophet secretary problem is effectively a distribution over instances of the standard prophet inequalities problem. We refer to Appendix A for more details.*

2 Related work

In this paper we primarily focus on prophet inequalities for problems with combinatorial feasibility constraints, where the economic objective to maximize is the social welfare. However, in the discussion below we also mention the economic objective of revenue in few contexts.

Dütting et al. [18, 19] raise as an open question whether any prophet inequality algorithm can be turned into pricing-based algorithm. They specifically ask whether there exist prophet inequalities which cannot be implemented using posted prices. Our main result answers the former question in positive, and the latter question in negative, whenever there exists an algorithm for a given prophet inequality. Dütting et al. [19] present a general technique to establish prophet inequalities based on posted prices. Their technique, however, requires to establish an existence of a so called balanced pricing for the given problem, which is inspired by the smoothness framework for the price of anarchy. Balanced pricing is problem specific and their technique does not allow to transform any prophet inequality algorithm into a pricing based one, unlike our technique.

Correa and Cristi [16] prove an existence of the first constant factor prophet inequality for online combinatorial auctions with subadditive valuations, maximizing the social welfare. Their result is purely existential, and they pose as an important open problem whether there exist posted prices implementing their algorithm. They also explicitly ask for an existence of a reduction that given any prophet inequality algorithm, transforms it into a posted pricing based algorithm with the same (or almost the same) approximation ratio. Our technique provides such general reduction and implies the existence of such posted prices, answering their open questions. Feldman et al. [25] obtained a prophet inequality of factor $\log(m)$ for CAs with subadditive valuations, where m is the number of items. This result was recently improved to a factor $\log\log(m)$ by Dütting et al. [20]. Feldman et al. [25] designed the best possible prophet inequality algorithm for combinatorial auctions with XOS, and therefore also submodular, valuations, which is also based on item posted pricing. Their algorithm achieves the approximation ratio of 2, which is best possible even for single item prophet inequality, see [36, 37]. The algorithms in [25, 36, 37] are based on anonymous item pricing.

When the arrival order of the agents is uniformly random, rather than fixed, Ehsani et al. [21] improve this approximation factor to the optimal factor of $1 - 1/e$ for combinatorial auctions with XOS valuations. They also prove the same factor $1 - 1/e$ when agents, arriving in a uniformly random order, correspond to elements of a matroid and the solution has to be an independent set of that matroid. Their algorithms are threshold/pricing-based.

Chawla et al. [9] pioneered the study of posted-price mechanisms in Bayesian mechanism design for various problems with combinatorial feasibility constraints. They design revenue maximizing DSIC mechanisms for problems that involve matroid (intersection) constraints and multi-unit multi-item unit-demand problems, that guarantee constant approximation of the optimal revenue. These guarantees also apply to the social welfare maximization for these problems. Many of these results have been significantly improved by Alaei [2] who has provided a reduction from multi-buyer to single-buyer setting for combinatorial auctions and related Bayesian mechanism design problems with combinatorial feasibility constraints.

Ezra et al. [23, 24] design algorithms for prophet inequality problems with matching constraints. They do not use a pricing based approach, but build on the technique of online contention resolution schemes (OCRS), extending it to what they call batched OCRS technique. Using this new technique for matching with vertex arrivals, they extend the result of [25] for bipartite graphs to general graphs and for matching with edge arrivals, they improve the $1/3$ approximation of [27] to 0.337. They pose as an open problem whether there exist pricing based algorithms with competitive ratios comparable to theirs. Our technique answers this open problem in affirmative, implying pricing based prophet inequality algorithms with the same approximation ratios.

We would also like to mention that Correa et al. [14] prove that designing posted price mechanisms (PPM) is equivalent to prophet inequalities based on thresholds. In fact they note that the implication from prophet inequality to PPM has been implicitly shown in [9, 31]. They prove the other implication from PPM to prophet inequality. However, their reduction works only for single-dimensional agents and they have focused on revenue maximization. While we focus on the social welfare, our reduction is much more general and stronger. Firstly, we do not assume that we have a threshold based prophet inequality algorithm, but any prophet inequality algorithm, and we show how to transform it into a PPM with the same performance guarantee. And secondly, we treat a very general setting with possibly multi-dimensional agents.

3 Preliminaries

Notation. Given an integer n, we use $[n]$ to denote the set $\{1, \ldots, n\}$. Given a vector $(a_1, \ldots, a_n)$, we use $a_{\leq i} := (a_1, \ldots, a_i)$ to denote the first i elements of the vector, and define $a_{<i} := a_{\leq i-1}$. We use $\Pr[\,.\,]$ and $\mathbb{E}[\,.\,]$ to denote probabilities and expectations respectively. For an event A, we define its indicator $\mathbb{1}\{A\}$ as the random variable that equals 1 if A holds and equals 0 otherwise. Note that $\Pr[A] = \mathbb{E}[\mathbb{1}\{A\}]$. For random variables X, Y, and an event A, we use $\mathbb{E}[X \mid Y]$ and $\mathbb{E}[X \mid A]$ to denote the conditional expectation of X with respect to Y and A respectively. For any set X, we use Δ_X to denote the set of all distributions over X.

Model. We consider a setting with n agents N where each agent $i \in N$ has an outcome space X_i. The joint outcome space is denoted by $\mathcal{X} := X_1 \times \cdots \times X_n$. and the set of feasible outcomes is denoted by $\mathcal{F} \subseteq \mathcal{X}$. Each agent also has a distribution D_i over the set of valuation functions V_i, where a valuation function $v_i \in V_i$ is a mapping $v_i : X_i \to \mathbb{R}^{\geq 0}$, describing the agent's valuation for each outcome.

The agents arrive sequentially, with some order $\sigma : [n] \to N$ that is not known to the algorithm. In order to keep the notation simple, we will denote the agent arriving at time i with i[1]. Upon arrival, each agent i reveals its identity and its valuation function $v_i \sim D_i$, and an algorithm for the problem needs to assign it to some outcome in X_i. Given an algorithm $\mathcal{A}$, and valuation functions $\mathbf{v} = (v_1, \ldots, v_n)$, we use $\mathcal{A}_i(\mathbf{v})$, to denote the outcome assigned to i when the algorithm observes agents with the valuations functions v_i, and use $\mathcal{A}_{\leq i}(\mathbf{v}) := (\mathcal{A}_1(\mathbf{v}), \ldots, \mathcal{A}_i(\mathbf{v}))$ to denote the set of outcomes assigned to the first i arriving agents. We also define $\mathcal{A}_{<i}(\mathbf{v}) := \mathcal{A}_{\leq i-1}(\mathbf{v})$. We will drop the dependence on v from the notation when it is clear from the context. We will also abuse notation and use $\mathcal{A}$ to denote the final assignment $\mathcal{A}_{\leq n}$. Note that even for fixed values of v, the outcomes $\mathcal{A}_i(\mathbf{v})$ can be random because of the algorithm's randomness. This effectively means that we can think of $\mathcal{A}$ as a randomized mapping from $V_1 \times \cdots \times V_n$ to $\mathcal{X}$. An algorithm is correct if its assigned outcome is always feasible, i.e., $\mathcal{A}_{\leq n} \in \mathcal{F}$.

We say an algorithm $\mathcal{A}$ is *pricing based* if for each agent i, it chooses a *pricing function* $\pi_i : X_i \to \mathbb{R}$, and assigns i to some outcome in $\arg\max_{x \in X_i} (\, v_i(x) - \pi(x) \,)$, where ties can be broken arbitrarily and possibly randomly.[2] The choice of the pricing function π_i can depend on the valuation distribution D_i, the previously observed valuations $v_1, \ldots v_{i-1}$, and the set of already assigned outcomes. Crucially however, the pricing function does not depend on v_i and needs to be chosen before observing v_i. In the language of algorithmic mechanism design, a pricing based algorithm corresponds to a *dominant strategy incentive compatible (DSIC)* mechanism with bidders given by *demand oracles/queries*, see, e.g., [6].

While in principle, an algorithm's decisions for agent i may depend on past valuation functions v_i, for most existing algorithms this is not the case. We say an algorithm is *past-valuation independent* if for determining the assignment to agent i it only uses the current valuation v_i, and the set of previously assigned outcomes $\mathcal{A}^{\text{out}}_{<i}$.

Given valuation functions $\mathbf{v} = (v_1, \ldots, v_n)$, we define the valuation of algorithm $\mathcal{A}$ as the social welfare of the agents, i.e.,

$$\mathbf{v}(\mathcal{A}) := \sum_i v_i(\mathcal{A}_i)\ .$$

In general, the objective of the prophet inequalities problem is to design an algorithm that maximizes $\mathbb{E}[\,\mathbf{v}(\mathcal{A})\,]$ while ensuring the constraint $\mathcal{A} \in \mathcal{F}$.

4 Black-box reduction

In this section, we present our black-box reduction. We first state our main result, which is a more formal restatement of Theorem 1 from the introduction.

THEOREM 4. *Let $\mathcal{A}^{\text{inp}}$ be an algorithm for the prophet inequalities problem, and let $\mathbb{E}\left[\,\mathbf{v}(\mathcal{A}^{\text{inp}})\,\right]$ denote its expected social welfare. There exists an algorithm $\mathcal{A}^{\text{out}}$ with the following properties:*

- ***Identical distribution.*** *The distribution of $\mathcal{A}^{\text{out}}$ and $\mathcal{A}^{\text{inp}}$ over the assigned outcomes are the same. Formally, letting $\mathcal{A}_i \in X_i$ denote the assignment made by algorithm $\mathcal{A}$ to agent i,*

$$\Pr\left[\,(\mathcal{A}^{\text{inp}}_1, \ldots, \mathcal{A}^{\text{inp}}_n) = (x_1, \ldots, x_n)\,\right] = \Pr\left[\,(\mathcal{A}^{\text{out}}_1, \ldots, \mathcal{A}^{\text{out}}_n) = (x_1, \ldots, x_n)\,\right] \tag{4.3}$$

[1] This does not contradict the assumption that the arrival order is a priori unknown because for each agent i, our algorithm will not use information about the arrival order of agents $i+1, \ldots, n$.

[2] In principle, the definition allows the tie-breaking distribution to depend on the valuation function v_i. The reduction we provide however has the additional property that the distribution depends only on the set of tied outcomes (see Lemma 10).

for all $(x_1, \ldots, x_n) \in \mathcal{X}$. *The randomness here comes from the randomness of the valuation functions* $v_1, \ldots, v_n$, *as well as the randomness in the algorithms* $\mathcal{A}^{\text{inp}}$ *and* $\mathcal{A}^{\text{out}}$.

- ***Non-decreasing welfare.*** *The expected social welfare under algorithm* $\mathcal{A}^{\text{out}}$ *is not less than the expected social welfare under* $\mathcal{A}^{\text{inp}}$. *Formally,* $\mathbb{E}[\mathbf{v}(\mathcal{A}^{\text{out}})] \geq \mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{inp}})\right]$. *Moreover, for each agent* i, $\mathbb{E}\left[v_i(\mathcal{A}_i^{\text{out}})\right] \geq \mathbb{E}\left[v_i(\mathcal{A}_i^{\text{inp}})\right]$.
- ***Pricing.*** *The algorithm* $\mathcal{A}^{\text{out}}$ *is pricing based.*

We note that for any fixed realizations of $v_1, \ldots, v_n$, the outcomes assigned by $\mathcal{A}^{\text{inp}}$ and $\mathcal{A}^{\text{out}}$ can be (and in general are) different. The identical distribution condition (4.3) simply states that the distribution of $\mathcal{A}^{\text{inp}}$ and $\mathcal{A}^{\text{out}}$ is the same. Intuitively, $\mathcal{A}^{\text{out}}$ redistributes the output distribution of $\mathcal{A}^{\text{inp}}$ for different valuation functions: For any vector $(x_1, \ldots, x_n) \in \mathcal{X}$, changing $\mathcal{A} = \mathcal{A}^{\text{inp}}$ to $\mathcal{A} = \mathcal{A}^{\text{out}}$ increases the probability $\Pr[\mathcal{A} = (x_1, \ldots, x_n) \mid v_1, \ldots, v_n]$ for some values of $v_1, \ldots, v_n$, and decreases the probability for other values, without changing the overall probability $\Pr[\mathcal{A} = (x_1, \ldots, x_n)]$.

We will present the algorithm $\mathcal{A}^{\text{out}}$ in Section 4.1. Our algorithm will be based on the notion of *assignment rules*, which we will define shortly, and will essentially implement an optimal assignment rule subject to some constraints. in Section 4.2, we show how this assignment rule can be be calculated, and prove that it leads to a pricing based algorithm. In Section 4.3, we prove Theorem 4.

4.1 Algorithm

Assignment rules. Before providing the algorithm $\mathcal{A}^{\text{out}}$, we define the concept of an *assignment rule*, which characterizes the decision making process for each individual agent. For any $i \in [n]$, we define an assignment rule $A_i : V_i \to \Delta_{X_i}$ as any mapping from the space of valuation functions V_i to the set of distributions over outcomes X_i. We will abuse notation and simply write $A_i(v) \in X_i$ to denote a random sample drawn from the mapped distribution. Given this definition, an algorithm for the problem is simply a sequence of assignment rules $(A_1, \ldots, A_n)$, where each A_i is determined based on the previously observed agents $[i-1]$, the identity of the i-th agent, and the valuation functions $v_1, \ldots, v_{i-1}$. We say an assignment rule A_i is *pricing based* if there exists a *pricing function* $\pi_i : X_i \to \mathbb{R}$, such that $A_i(v) \in \arg\max_{x \in X_i} (v_i(x) - \pi(x))$ for all i, x with probability 1. Ties can be broken arbitrarily and possibly randomly. It is clear that an algorithm $\mathcal{A}$ is pricing based if and only if its corresponding assignment rules $A_1, \ldots, A_n$ are guaranteed to be pricing based.

The assignment rule view of the algorithm is important conceptually as it separates the decision making process for any algorithm $\mathcal{A}$ to two parts. First, $\mathcal{A}$ commits to an assignment rule $A_i : V_i \to \Delta_{X_i}$ based on everything it has observed before seeing v_i. It then sees v_i and samples the outcome $A_i(v_i)$.

Analysis of a single agent. We now explain our algorithm by considering a fixed agent $i \in [n]$, and describing the assignment rule for this agent. Before the arrival of agent i, the algorithm has observed the arrival order of the first $i-1$ agents, the valuation functions $v_1, \ldots, v_{i-1}$, and has assigned them the outcome $\mathcal{A}^{\text{out}}_{<i}$. These random variables have therefore been realized at this point and, since we are focusing on a single agent i, we will assume that $v_1, \ldots, v_{i-1}, \mathcal{A}^{\text{out}}_{<i}$ are fixed, deterministic values. When considering the online sequence as a whole, i.e., not just a single i, the claims we make in this section will hold conditioned on $v_1, \ldots, v_{i-1}, \mathcal{A}^{\text{out}}_{<i}$ (see Remark 5).

As mentioned in the introduction, on a high level, we first consider the assignments of the input algorithm $\mathcal{A}^{\text{inp}}$, and then improve upon this without changing the assignment distribution. For any set of outcomes $x_{<i} = (x_1, \ldots, x_{i-1}) \in X_1 \times X_2 \times \ldots \times X_{i-1}$, we define $\overline{A}^{\text{inp}}_{x_{<i}}$ as the "averaged" version of the assignment rule A_i^{inp} corresponding to $\mathcal{A}^{\text{inp}}$, when $\mathcal{A}^{\text{inp}}_{<i} = x_{<i}$. Formally, the assignment rule $\overline{A}^{\text{inp}}_{x_{<i}} : V_i \to X_i$ is defined via the following distribution over X_i for each $v \in V_i$:[3]

$$\Pr\left[\overline{A}^{\text{inp}}_{x_{<i}}(v) = x\right] := \Pr_{v_1' \sim D_1, \ldots, v_{i-1}' \sim D_{i-1}}\left[\mathcal{A}_i^{\text{inp}}(v_1', \ldots, v_{i-1}', v) = x \mid \mathcal{A}^{\text{inp}}_{<i}(v_1', \ldots, v_i') = x_{<i}\right], \tag{4.4}$$

In the above definition, the randomness is over the randomness of the draws of $v_1', \ldots, v_i'$, as well as the internal randomness of algorithm $\mathcal{A}^{\text{inp}}$.[4] Intuitively, the definition captures what the input algorithm *would do on average* if it had assigned the outcomes $x_{<i}$ before the arrival of agent i. We note that in general, the behaviour of algorithm $\mathcal{A}^{\text{inp}}$ may depend on the valuation functions $v_1, \ldots, v_{i-1}$. The above definition averages out this dependence however, and our definition only conditions on the value of $\mathcal{A}^{\text{inp}}_{<i}$. This will be important in the design of our algorithm as we will preserve the distribution of the input algorithm over outcomes (see the identical distribution condition in Theorem 4), but we will not necessarily make the same assignments as the input algorithm for any individual v_i.

[3]As we will see, $\Pr\left[\mathcal{A}^{\text{inp}}_{<i} = x_{<i}\right]$ will always be strictly positive for all $x_{<i}$ in our algorithm, so the conditional expectation is always well-defined.

[4]For simplicity, throughout the paper we omit the dependence on the internal randomness of our algorithms in the notation; e.g., we write $\Pr_{v_1, \ldots, v_i}[]$ instead of $\Pr_{v_1, \ldots, v_i, \mathcal{A}_i^{\text{inp}}}[]$.

Define $\overline{A}_i^{\text{inp}} := \overline{A}^{\text{inp}}_{\mathcal{A}^{\text{out}}_{<i}}$. We note that since $\mathcal{A}^{\text{out}}_{<i}$ was assumed to be fixed, $\overline{A}_i^{\text{inp}}$ is fixed as well. Our algorithm will find an assignment rule A_i^{out} that improves $\overline{A}_i^{\text{inp}}$, while preserving its distribution for $v_i \sim D_i$. Specifically, the assignment rule will satisfy the three properties stated below, which correspond to the properties of $\mathcal{A}^{\text{out}}$ as described in Theorem 4.

1. **Identical distribution.** $\Pr_{v\sim D_i}\left[A_i^{\text{out}}(v) = x_i\right] = \Pr_{v\sim D_i}\left[\overline{A}_i^{\text{inp}}(v) = x_i\right]$ for all $x_i \in X_i$,
2. **Non-decreasing welfare.** $\mathbb{E}_{v\sim D_i}\left[v(A_i^{\text{out}}(v))\right] \geq \mathbb{E}_{v\sim D_i}\left[v(\overline{A}_i^{\text{inp}}(v))\right]$,
3. **Pricing.** A_i^{out} is pricing based, with a corresponding pricing function π_i.

Remark 5. *Since we focused on a single agent, we stated our results for fixed values of $v_1, \ldots, v_{i-1}, \mathcal{A}^{\text{out}}_{<i}$ and the randomness in the above properties is over the draw of v and the randomness of the assignment rules A_i^{out} and $\overline{A}_i^{\text{inp}}$. When considering the algorithm for all agents however, $v_1, \ldots, v_{i-1}, \mathcal{A}^{\text{out}}_{<i}$ will be random variables and the stated properties will hold conditioned on them. Specifically,*

$$\Pr\left[A_i^{\text{out}}(v_i) = x \mid v_{<i}, \mathcal{A}^{\text{out}}_{<i}\right] = \Pr\left[\overline{A}_i^{\text{inp}}(v_i) = x \mid v_{<i}, \mathcal{A}^{\text{out}}_{<i}\right],$$

and

$$\mathbb{E}\left[v_i(A_i^{\text{out}}(v_i)) \mid v_{<i}, \mathcal{A}^{\text{out}}_{<i}\right] \geq \mathbb{E}\left[v_i(\overline{A}_i^{\text{inp}}(v_i)) \mid v_{<i}, \mathcal{A}^{\text{out}}_{<i}\right].$$

Implementation of assignment rules. In order to implement an assignment rule with the properties specified above, we find the assignment rule that maximizes social welfare, subject to the identical distribution constraint. We will show that this naturally leads to a pricing based solution. As before, since we are focusing on a single agent, we will assume that $v_1, \ldots, v_{i-1}, \mathcal{A}^{\text{out}}_{<i}$ (and by extension $\overline{A}_i^{\text{inp}}$) are fixed deterministic values.

We first calculate the probability distribution $\Pr_{v\sim D_i}\left[\overline{A}_i^{\text{inp}}(v) = x_i\right]$. For any algorithm $\mathcal{A}$, and any $x_{<i}, x_i$, we define the *$\mathcal{A}$-likelihood of x_i given $x_{<i}$*, denoted by $p_{\mathcal{A}}(x_i; x_{<i})$, as

$$p_{\mathcal{A}}(x_i; x_{<i}) := \Pr_{v_1'\sim D_1,\ldots,v_i'\sim D_i}\left[\mathcal{A}_i(\mathbf{v}') = x_i \mid \mathcal{A}_{<i}(\mathbf{v}') = x_{<i}\right]. \tag{4.5}$$

We further define $p_i(x) := p_{\mathcal{A}^{\text{inp}}}(x; \mathcal{A}^{\text{out}}_{<i})$. Note that $p_i(x)$ is a fixed value for any x as we assumed that $\mathcal{A}^{\text{out}}_{<i}$ was fixed. While we do not focus on the computational aspects in this section, we note that the above expression can be calculated using Monte-Carlo simulation up to arbitrary precision.

It can be shown (see Lemma 12) that $p_i(x) = \Pr_{v\sim D_i}\left[\overline{A}_i^{\text{inp}}(v) = x\right]$. We therefore formulate the problem of finding the assignment rule with maximum social welfare as the following optimization problem over all assignment rules $A : V \to \Delta_X$, with (X, V, D, p) set to (X_i, V_i, D_i, p_i):

$$\begin{aligned} \max_A \quad & \mathbb{E}_{v\sim D}\left[v(A(v))\right] \\ \text{s.t.} \quad & \Pr_{v\sim D}\left[A(v) = x\right] = p(x) \text{ for all } x \in X\,. \end{aligned} \tag{P1}$$

The solution to this optimization problem, which we denote by A_i^{out}, satisfies the identical distribution condition because of the optimization problem's constraints. Additionally, since $\overline{A}_i^{\text{inp}}$ is feasible for the optimization problem (P1) and A_i^{out} is the optimal solution, we conclude that $\mathbb{E}\left[v_i(A_i^{\text{out}}(v_i))\right] \geq \mathbb{E}\left[v_i(\overline{A}_i^{\text{inp}}(v_i))\right]$, which means that A_i^{out} satisfies the non-decreasing welfare condition as well. We will show that the solution to this optimization problem is a pricing based assignment rule, which will ensure the pricing based condition. The condition follows from the lemma below, the proof of which is in Section 4.2.

Lemma 6. *There is an optimal solution to the optimization problem* (P1) *that is pricing based. Moreover, any pricing based assignment rule satisfying the constraints of the optimization problem is an optimal solution.*

A pseudocode of our approach is given in Algorithm 1. We note that we do not need to calculate $\overline{A}_i^{\text{inp}}$ explicitly in our algorithm, and only require p_i.

4.2 Analysis of the optimal assignment problem In order to prove Lemma 6, we use a duality argument. We first provide an alternative characterization of assignment rules as a joint distribution over the space of valuation functions and the assigned

Algorithm 1: Black-box reduction

Input: Input parameters $n, D_1, \dots, D_n, \mathcal{A}^{\text{inp}}$
Output: Assignment outcomes $\mathcal{A}^{\text{out}}_1, \dots, \mathcal{A}^{\text{out}}_n$
for $i \in [n]$ **do**
  Calculate $p_i(x) = \Pr_{v'_1\sim D_1,\dots,v'_i\sim D_i}\left[\mathcal{A}^{\text{inp}}_i(\mathbf{v}') = x \mid \mathcal{A}^{\text{inp}}_{<i}(\mathbf{v}') = \mathcal{A}^{\text{out}}_{<i}\right]$ for $x \in X$;
  Solve the optimization problem (P1) with (D_i, p_i), obtaining the solution A^{out}_i;
  Observe the valuation function of the i-th agent v_i;
  Draw $\mathcal{A}^{\text{out}}_i$ from $A^{\text{out}}_i(v_i)$ and assign it to agent i;
end

outcomes. This characterization leads to a reformulation of the optimization problem (P1) as a linear program. Analysing the dual of this linear program then implies that its optimal solution corresponds to a pricing based algorithm. Throughout the section, we omit the dependence on i in the notation (e.g., write A instead of A_i) as we are considering a fixed i.

We start with the following definition.

Definition 7. (Associated coupling) *Given an assignment rule A mapping V to X under D, we define its* associated coupling *$\gamma_{A,D}$ as as the distribution of $(v, A(v))$ where v is sampled from D. Note that $\gamma_{A,D}$ is a joint distribution over $V \times X$. We omit the dependence on D when it is clear from context.*

The above definition allows us to work with joint distributions instead of assignment rules. Formally, we prove the following lemma in the Appendix.

Lemma 8. *For an assignment rule $A : V \to \Delta_X$, and a distribution D, let $A(D)$ denote the distribution of $A(v)$ over X for $v \sim D$, i.e., $A(D)(x) = \Pr_{v\sim D}[A(v) = x]$. The marginal of $\gamma_{A,D}$ on V and X equals D and $A(D)$ respectively. Additionally, $\mathbb{E}_{v\sim D}[v(A(v))] = \mathbb{E}_{(v,x)\sim\gamma_{A,D}}[v(x)]$*

The notion of a pricing based assignment rule can easily be extended to associated coupling as well.

Definition 9. (Pricing-based coupling) *The joint distribution γ over $V \times X$ is called pricing based if there exist functions $\pi : X \to \mathbb{R} \cup \{+\infty\}$ and $\psi : V \to \mathbb{R} \cup \{+\infty\}$ such that*

1. *$\psi(v) + \pi(x) \geq v(x)$ for all v, x, and*
2. *The probability of the set $\{(v, x) : \psi(v) + \pi(x) > v(x)\}$ is zero under γ.*

Pricing based assignment rules correspond to pricing based couplings, as we show in the lemma below, the proof of which is in the Appendix.

Lemma 10. *If A is a pricing based assignment rule with pricing function π, then $\gamma_{A,D}$ is a pricing based coupling with the pricing function π. Conversely, if γ is a pricing based coupling with the pricing function π, and its marginal on V equals D, there exists an assignment rule A with pricing function π such $A(D)$ equals the marginal of γ on X, and $\mathbb{E}_{(v,x)\sim\gamma}[v(x)] = \mathbb{E}_{(v,x)\sim\gamma_{A,D}}[v(x)]$.*

Given the above lemmas, the optimization problem (P1) can be rewritten as the following linear program:

$$\begin{aligned} \max_{\gamma} \quad & \mathbb{E}_{(v,x)\sim\gamma}[v(x)] \\ \text{s.t.} \quad & \gamma \in \Gamma(D, p)\ , \end{aligned} \tag{LP1}$$

where $\Gamma(D, p)$ denotes the set of all distributions over $V \times X$ with marginals D on V and p on X. The optimization problem (LP1) is an *optimal transport* problem which has applications in mathematics (see Villani [43] for a detailed overview) and economics [42, 28, 11]. Similar, in spirit, duality approaches were used, for example, in context of combinatorial auctions and Walrasian equilibria, see, e.g., [6]. Unlike these works, however, we use the optimization problem only for a single agent to find the best personalized assignment. Whereas, the mentioned applications study the whole population of agents and find an assignment of items to each agent, and corresponding prices for the items, that optimizes social welfare / clears the market.

Lemma 11. *Assume that the optimization problem* (LP1) *is feasible. The joint distribution $\gamma \in \Gamma(D, p)$ is the optimal solution to* (LP1) *if and only if it satisfies the constraint and it is a pricing based coupling.*

Proof. We provide a proof here for discrete D as it provides a nice intuition. The proof for the general case follows from Kantorovich duality; specifically, we refer to Villani [43], Theorem 5.10, part (ii), the equivalence of conditions (a) and (c).

The optimization problem (LP1) can be rewritten as:

$$\min_{\gamma} \sum_{v,x} -v(x)\gamma(v,x)$$
$$\forall v : \sum_{x} \gamma(v,x) = D(v)$$
$$\forall x : \sum_{v} \gamma(v,x) = p(x)$$
$$\forall v,x : \gamma(v,x) \geq 0 \ .$$

Since the linear program is feasible, strong duality holds and we can use the KKT conditions to characterize the problem's solution. Forming the Lagrangian, we obtain

$$\sum_{v,x} -v(x)\gamma(v,x) + \sum_{v,x} \psi(v)\gamma(v,x) - \sum_{v} \psi(v)D(v) + \sum_{v,x} \pi(x)\gamma(v,x) - \sum_{x} \pi(x)p(x) - \sum_{v,x} \rho(v,x)\gamma(v,x) \ .$$

The dual optimality conditions can therefore be written as

$$-v(x) + \psi(v) + \pi(x) - \rho(v,x) = 0 \ ,$$

which, given the dual constraint $\rho(v,x) \geq 0$ is equivalent to

$$\psi(v) + \pi(x) \geq v(x).$$

The dual objective equals $\max\left(-\sum \psi(v)D(v) - \sum \pi(x)p(x)\right)$, which is equivalent to

$$\min \sum_{v} \psi(v)D(v) + \sum_{x} \pi(x)p(x) \ .$$

Therefore, the dual problem can be written as

$$\min \sum_{v} \psi(v)D(v) + \sum_{x} \pi(x)p(x)$$
$$\forall v,x : \psi(v) + \pi(x) \geq v(x) \ .$$

The complementary slackness implies that whenever $\gamma(v,x) > 0$, we should have $\psi(v) + \pi(x) = v(x)$, which means that γ is pricing based. Note that this means that for any x such that $\gamma(v,x) > 0$ we have $x \in \arg\max_{x'}\left(v(x') - \pi(x')\right)$ because

$$v(x) - \pi(x) = \psi(v) \geq v(x') - \pi(x')$$

for all $x' \in X$. Therefore, if γ is an optimal solution, then it must be pricing based.

Conversely, if γ is a pricing based solution satisfying the primal feasibility, then it satisfies the dual feasibility and complementary slackness condition because it is pricing based. Therefore, it satisfies the KKT condition which means it is an optimal solution. □

We can now prove Lemma 6.

Proof. [Proof of Lemma 6] By Lemma 12, $\overline{A}_i^{\text{inp}}$ is feasible for (P1). By Lemma 8, this implies that $\gamma_{\overline{A}_i^{\text{inp}}}$ is feasible for (LP1). Let γ be the solution to the optimization problem (LP1), which we assume is pricing based given Lemma 11. Let A be the corresponding pricing based algorithm, as specified by Lemma 10. The assignment rule A satisfies the constraint of (P1) given Lemma 10 and

$$\mathbb{E}_{v\sim D}\left[v(A(v))\right] = \mathbb{E}_{(v,x)\sim\gamma}\left[v(x)\right] \ .$$

Furthermore, for any assignment rule A' satisfying the constraints, $\gamma_{A'}$ satisfies the constraint of (LP1) which implies

$$\mathbb{E}_{v\sim D}\left[v(A'(v))\right] = \mathbb{E}_{(v,x)\sim\gamma_{A'}}\left[v(x)\right] \leq \mathbb{E}_{(v,x)\sim\gamma}\left[v(x)\right] \ ,$$

where the equality follows from Lemma 8, and the inequality follows from the optimality of γ. Therefore, $\mathbb{E}_{v\sim D}\left[v(A(v))\right] \geq \mathbb{E}_{v\sim D}\left[v(A'(v))\right]$, finishing the proof. □

4.3 Proof of Theorem 4 We now prove Theorem 4.

Overview of the proof. We here briefly explain the role of each of the three conditions for assignment rule, and how together they prove Theorem 4. Expressed in terms of $\mathcal{A}^{\text{out}}$, the identical distribution condition simply states that $p_{\mathcal{A}^{\text{out}}}(x_i; x_{<i}) = p_{\mathcal{A}^{\text{inp}}}(x_i; x_{<i})$ for all $x_{<i}, x_i$ (see Lemma 13). In other words, averaged over the randomness of v_i, the conditional distribution of $\mathcal{A}_i^{\text{out}}$ given $\mathcal{A}_{<i}^{\text{out}}$ is the same as the conditional distribution of $\mathcal{A}_i^{\text{inp}}$ given $\mathcal{A}_{<i}^{\text{inp}}$. Since $\mathcal{A}_0^{\text{out}} = \mathcal{A}_0^{\text{inp}} = \emptyset$, an inductive argument shows that $\mathcal{A}^{\text{out}}$ and $\mathcal{A}^{\text{inp}}$ have the same *joint* distribution over the outcome space (Lemma 14). This gives the identical distribution condition required by Theorem 4, i.e., Equation (4.3).

Equation (4.3) has two important consequences. Firstly, whenever the assignment of $\mathcal{A}^{\text{inp}}$ is feasible for some combinatorial constraint, it implies that the assignment of $\mathcal{A}^{\text{out}}$ is always feasible as well. More importantly however, it ensures that we can compare the expected welfare of $\mathcal{A}^{\text{out}}$ and $\mathcal{A}^{\text{inp}}$ by comparing the expected valuation of A_i^{out} and $\overline{A}_i^{\text{inp}}$ for each i, which is done using the non-decreasing welfare condition. Together with (4.3), the non-decreasing welfare condition implies that $\mathbb{E}\left[\, v_i(\mathcal{A}_i^{\text{out}})\,\right] \geq \mathbb{E}\left[\, v_i(\mathcal{A}_i^{\text{inp}})\,\right]$ for all i (Lemma 16), which implies $\mathbb{E}\left[\,\mathbf{v}(\mathcal{A}^{\text{out}})\,\right] \geq \mathbb{E}\left[\,\mathbf{v}(\mathcal{A}^{\text{inp}})\,\right]$. Finally, the pricing condition for the assignment rules A_i^{out} implies that the algorithm $\mathcal{A}^{\text{out}}$ is pricing based as required by the theorem.

Lemma 12. *For any i and $(x_1, \ldots, x_i) \in X_1 \times \ldots X_i$,*

$$\Pr_{v\sim D_i}\left[\overline{A}_{x_{<i}}^{\text{inp}}(v) = x_i\right] = p_{\mathcal{A}^{\text{inp}}}(x_i; x_{<i})\,.$$

Proof. For any fixed value of v, by definition of $\overline{A}_{x_{<i}}^{\text{inp}}$,

$$\Pr\left[\overline{A}_{x_{<i}}^{\text{inp}}(v) = x\right] = \Pr_{v_1'\sim D_1,\ldots,v_{i-1}'\sim D_{i-1}}\left[A_i^{\text{inp}}(v_1', \ldots, v_{i-1}', v) = x \mid \mathcal{A}_{<i}^{\text{inp}}(v_{\leq i}') = x_{<i}\right].$$

Therefore, by iterated expectation,

$$\Pr_{v_i'\sim D_i}\left[\overline{A}_{x_{<i}}^{\text{inp}}(v_i') = x\right] = \Pr_{v_1',\ldots,v_i'}\left[\mathcal{A}_i^{\text{inp}}(\mathbf{v}') = x \mid \mathcal{A}_{<i}^{\text{inp}}(\mathbf{v}') = x_{<i}\right],$$

which equals $p_{\mathcal{A}^{\text{inp}}}(x; x_{<i})$. ∎

Lemma 13. *Let $\mathcal{A}^{\text{out}}$ be an algorithm such that for any agent i, its assignment rule A_i satisfies the identical distribution condition. For any $(x_1, \ldots, x_i)$, we have $p_{\mathcal{A}^{\text{out}}}(x_i; x_{<i}) = p_{\mathcal{A}^{\text{inp}}}(x_i; x_{<i})$.*

Proof. For any $(x_1, \ldots, x_i)$,

$$\begin{aligned}
\Pr_{v_1,\ldots,v_i}\left[\overline{A}_i^{\text{inp}}(v_i) = x_i \mid \mathcal{A}_{<i}^{\text{out}} = x_{<i}\right] &= \Pr_{v_1,\ldots,v_i}\left[\overline{A}_{x_{<i}}^{\text{inp}}(v_i) = x_i \mid \mathcal{A}_{<i}^{\text{out}} = x_{<i}\right] && \textit{(Definition of } \overline{A}_i^{\text{inp}}\textit{)}\\
&= p_{\mathcal{A}^{\text{inp}}}(x_i; x_{<i})\,. && \textit{(Lemma 12)}
\end{aligned} \tag{4.6}$$

By definition of $\mathcal{A}^{\text{out}}$, we have $\mathcal{A}_i^{\text{out}} = A_i^{\text{out}}(v_i)$. Therefore,

$$\begin{aligned}
p_{\mathcal{A}^{\text{out}}}(x_i; x_{<i}) &= \Pr_{v_1\sim D_1,\ldots,v_i\sim D_i}\left[\mathcal{A}_i^{\text{out}} = x_i \mid \mathcal{A}_{<i}^{\text{out}} = x_{<i}\right] && \textit{(Definition of } p_{\mathcal{A}^{\text{out}}}\textit{)}\\
&= \Pr_{v_1\sim D_1,\ldots,v_i\sim D_i}\left[A_i^{\text{out}}(v_i) = x_i \mid \mathcal{A}_{<i}^{\text{out}} = x_{<i}\right] && \textit{(Definition of } \mathcal{A}_i^{\text{out}}\textit{)}\\
&= \Pr_{v_1\sim D_1,\ldots,v_i\sim D_i}\left[\overline{A}_i^{\text{inp}}(v_i) = x_i \mid \mathcal{A}_{<i}^{\text{out}} = x_{<i}\right] && \textit{(Remark 5)}\\
&= p_{\mathcal{A}^{\text{inp}}}(x_i; x_{<i})\,. && \textit{(Equation (4.6))}
\end{aligned}$$

∎

Lemma 14. *Let $\mathcal{A}^{\text{out}}$ be an algorithm such that for any agent i, its assignment rule A_i satisfies the identical distribution condition. Then for any $i \in [n]$,*

$$\forall x_{\leq i} : \Pr\left[\,\mathcal{A}_{\leq i}^{\text{out}} = x_{\leq i}\,\right] = \Pr\left[\,\mathcal{A}_{\leq i}^{\text{inp}} = x_{\leq i}\,\right].$$

Proof. We prove the claim using induction on t. For $t = 0$, the claim holds trivially as $\mathcal{A}_0^{\text{out}} = \mathcal{A}_0^{\text{inp}} = \emptyset$. Assuming the claim holds for $i-1$,

$$\begin{aligned}
\Pr\left[\mathcal{A}^{\text{out}}_{\le i}=x_{\le i}\right] &= \Pr\left[\mathcal{A}^{\text{out}}_{i}=x_i, \mathcal{A}^{\text{out}}_{<i}=x_{<i}\right] \\
&= \Pr\left[\mathcal{A}^{\text{out}}_{i}=x_i \mid \mathcal{A}^{\text{out}}_{<i}=x_{<i}\right]\Pr\left[\mathcal{A}^{\text{out}}_{<i}=x_{<i}\right] \\
&= \Pr\left[\mathcal{A}^{\text{out}}_{i}=x_i \mid \mathcal{A}^{\text{out}}_{<i}=x_{<i}\right]\Pr\left[\mathcal{A}^{\text{inp}}_{<i}=x_{<i}\right] && \textit{(Induction hypothesis)} \\
&= \Pr\left[\mathcal{A}^{\text{inp}}_{i}=x_i \mid \mathcal{A}^{\text{inp}}_{<i}=x_{<i}\right]\Pr\left[\mathcal{A}^{\text{inp}}_{<i}=x_{<i}\right] && \textit{(Lemma 13)} \\
&= \Pr\left[\mathcal{A}^{\text{inp}}_{i}=x_i, \mathcal{A}^{\text{inp}}_{<i}=x_{<i}\right] \\
&= \Pr\left[\mathcal{A}^{\text{inp}}_{\le i}=x_{\le i}\right].
\end{aligned}$$

□

Lemma 15. *For any $t\in[n]$, and any fixed realization of $A^{\text{out}}_{<t}$, the expected valuation of agent t under $\mathcal{A}^{\text{inp}}$ conditioned on $\mathcal{A}^{\text{out}}_{<t}=A^{\text{out}}_{<t}$ is the same as the expected valuation of $\overline{A}^{\text{inp}}_t$. Formally, for all $x_{<t}$,*

$$\mathbb{E}_{v_t\sim D_t}\left[v_t(\overline{A}^{\text{inp}}_{x_{<t}}(v_t))\right]=\mathbb{E}_{v_1,\dots,v_t}\left[v_t(\mathcal{A}^{\text{inp}}_t)\mid \mathcal{A}^{\text{inp}}_{<t}=x_{<t}\right].$$

Proof. For any fixed value of v and any x, by definition of $\overline{A}^{\text{inp}}_{x_{<t}}$ (Equation (4.4)),

$$\Pr\left[\overline{A}^{\text{inp}}_{x_{<t}}(v)=x\right]=\Pr_{v_1',\dots,v_{t-1}'}\left[\mathcal{A}^{\text{inp}}_t(v_1',\dots,v_{t-1}',v)=x\mid \mathcal{A}^{\text{inp}}_{<t}(v_1',\dots,v_{t-1}')=x_{<t}\right],$$

where $v_i'\sim D_i$. It follows that, for any fixed v, considering only the randomness in $\overline{A}^{\text{inp}}_{x_{<t}}$,

$$\mathbb{E}\left[v(\overline{A}^{\text{inp}}_{x_{<t}}(v))\right]=\mathbb{E}_{v_1',\dots,v_{t-1}'}\left[v(\mathcal{A}^{\text{inp}}_t(v_1',\dots,v_{t-1}',v))\mid \mathcal{A}^{\text{inp}}_{<t}(v_1',\dots,v_{t-1}')=x_{<t}\right].$$

Setting v to $v_t'\sim D_t$ and taking expectation we obtain

$$\mathbb{E}_{v_t'\sim D_t}\left[v_t(\overline{A}^{\text{inp}}_t(v_t))\mid \mathcal{A}^{\text{out}}_{<t}=x_{<t}\right]=\mathbb{E}_{v_1',\dots,v_t'}\left[v_t(\mathcal{A}^{\text{inp}}_t(\mathbf{v}'))\mid \mathcal{A}^{\text{inp}}_{<t}(\mathbf{v}')=x_{<t}\right],$$

which proves the claim. □

Lemma 16. *Let $\mathcal{A}^{\text{out}}$ be an algorithm such that for any agent t, its assignment rule A_t satisfies the identical distribution and non-decreasing welfare condition. Then $\mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{out}}_t)\right]\ge\mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{inp}}_t)\right]$.*

Proof. Let $x_{<t}$ be a random sample from the distribution of $\mathcal{A}^{\text{out}}_{<t}$.

$$\begin{aligned}
\mathbb{E}\left[v_t(\mathcal{A}^{\text{out}}_t)\right] &= \mathbb{E}_{x_{<t}\sim\mathcal{A}^{\text{out}}_{<t}}\left[\mathbb{E}_{v_1,\dots,v_t}\left[v_t(\mathcal{A}^{\text{out}}_t)\mid \mathcal{A}^{\text{out}}_{<t}=x_{<t}\right]\right] && \textit{(Iterated expectation)} \\
&= \mathbb{E}_{x_{<t}\sim\mathcal{A}^{\text{out}}_{<t}}\left[\mathbb{E}_{v_1,\dots,v_t}\left[v_t(A^{\text{out}}_t(v_t))\mid \mathcal{A}^{\text{out}}_{<t}=x_{<t}\right]\right] && \textit{(Defintion of } \mathcal{A}^{\text{out}}_t\textit{)} \\
&\ge \mathbb{E}_{x_{<t}\sim\mathcal{A}^{\text{out}}_{<t}}\left[\mathbb{E}_{v_1,\dots,v_t}\left[v_t(\overline{A}^{\text{inp}}_t(v_t))\mid \mathcal{A}^{\text{out}}_{<t}=x_{<t}\right]\right], && \textit{(Non-decreasing welfare)}
\end{aligned} \tag{4.7}$$

By Lemma 15 however,

$$\begin{aligned}
\mathbb{E}_{v_1,\dots,v_t}\left[v_t(\overline{A}^{\text{inp}}_t(v_t))\mid \mathcal{A}^{\text{out}}_{<t}=x_{<t}\right] &= \mathbb{E}_{v_1,\dots,v_t}\left[v_t(\overline{A}^{\text{inp}}_{x_{<t}}(v_t))\mid \mathcal{A}^{\text{out}}_{<t}=x_{<t}\right] && \textit{(Definition of } \overline{A}^{\text{inp}}_t\textit{)} \\
&= \mathbb{E}_{v_t}\left[v_t(\overline{A}^{\text{inp}}_{x_{<t}}(v_t))\right] \\
&= \mathbb{E}_{v_1,\dots,v_t}\left[v_t(\mathcal{A}^{\text{inp}}_t)\mid \mathcal{A}^{\text{inp}}_t=x_{<t}\right]. && \textit{(Lemma 15)}
\end{aligned}$$

Plugging this back in Equation (4.7) implies

$$\begin{aligned}
\mathbb{E}\left[v_t(\mathcal{A}^{\text{out}}_t)\right] &\ge \mathbb{E}_{x_{<t}\sim\mathcal{A}^{\text{out}}_{<t}}\left[\mathbb{E}\left[v_t(\mathcal{A}^{\text{inp}}_t)\mid \mathcal{A}^{\text{inp}}_t=x_{<t}\right]\right] \\
&= \mathbb{E}_{x_{<t}\sim\mathcal{A}^{\text{inp}}_{<t}}\left[\mathbb{E}\left[v_t(\mathcal{A}^{\text{inp}}_t)\mid \mathcal{A}^{\text{inp}}_t=x_{<t}\right]\right] \\
&= \mathbb{E}\left[v_t(\mathcal{A}^{\text{inp}}_t)\right],
\end{aligned}$$

where the second equality follows from the fact that $\mathcal{A}^{\text{inp}}_{<t}$ and $\mathcal{A}^{\text{inp}}_{<t}$ have the same distribution (Lemma 14), and the third equality follows follows from iterated expectation. Summing over $t\in[n]$ finishes the proof. □

Proof. [Proof of Theorem 4] We use Algorithm 1. The identical distribution condition follows from Lemma 8, the non-decreasing welfare condition follows from Lemma 16, and the pricing condition follows from Lemma 6. ∎

5 Applications

Prophet inequalities for combinatorial auctions. In an online *Combinatorial Auction (CA)* we are given a set M of m items and a set $N = [n]$ of n agents. Each agent $i \in N$ has a valuation function $v_i : 2^M \to \mathbb{R}^{\geq 0}$, which is randomly and independently sampled from a given distribution D_i, defined over a set V_i of possible valuation functions. We assume that each possible realization of each v_i is monotone, i.e., $\forall A, B \subseteq M : A \subseteq B \Rightarrow v_i(A) \leq v_i(B)$. Agents arrive sequentially, w.l.o.g., in the order $1, 2, \ldots, n$, and upon arrival their valuation function is realized. At the time agent i arrives, the online algorithm has to decide which set $A_i \subseteq M \setminus \bigcup_{j=1}^{i-1} A_j$ of available items to allocate to agent i. This decision is irrevocable and the goal of the online algorithm is to maximize the *social welfare* of the allocation, i.e., the sum of the agents' valuations $\sum_{i \in N} v_i(A_i)$. A valuation function $v : 2^M \to \mathbb{R}^{\geq 0}$ is called *subadditive* if $\forall A, B \subseteq M : v(A \cup B) \leq v(A) + v(B)$.

In a recent breakthrough result, Correa and Cristi [16] prove an existence of a prophet inequality for online combinatorial auctions with subadditive valuations, which achieves an approximation factor of $6 + \varepsilon$, for any constant $\varepsilon > 0$, with respect to social welfare. This is the first known constant factor prophet inequality for this problem. However, their result is purely existential, where the existence of their algorithm is proved via a fixed point theorem of Kakutani. They also proved an existence of an incentive compatible implementation of their prophet inequality algorithm by using the same fixed point argument. This implementation is only a Bayesian Nash Equilibrium and not incentive compatible in dominant strategies (DSIC). If the arrival order of the agents is fixed and given, they also provided a dynamic programming implementation of the optimal online algorithm for this problem that implies existence of personalized posted bundle prices, and is thus a DSIC implementation. This optimal algorithm is highly inefficient but provides the same guarantee on the social welfare as their prophet inequality algorithm, because its performance is better than any online algorithm. Crucially, they pose as an important open problem whether there exist posted prices implementing their prophet inequality algorithm. They explicitly ask for an existence of a reduction that given any prophet inequality algorithm, transforms it into a posted pricing based algorithm with the same (or almost the same) approximation ratio. Our technique indeed provides such a general reduction and implies the existence of such posted prices, thus answering their open questions in the affirmative. This leads to the following result.

Theorem 17. *There exists an online posted pricing-based prophet inequality algorithm for social welfare maximizing combinatorial auctions with subadditive valuations with an expected approximation ratio of $\frac{1}{6}$. This algorithm implies a DSIC mechanism for this Bayesian mechanism design problem with the same approximation ratio.*

Proof. Let $\mathcal{A}$ be the the prophet inequality algorithm whose existence is proved in paper [16]. To see that the claim holds it suffices to show how to cast the problem of social welfare maximizing CAs with subadditive valuations in our general setting. Given agent $i \in N$, its outcome space is $X_i = 2^M$. The set of feasible outcomes $\mathcal{F} \subseteq \mathcal{X} = X_1 \times \cdots \times X_n$ is defined as $\mathcal{F} = \{(A_1, \ldots, A_n) \in (2^M)^n \mid \forall j, k \in [n] \,:\, j \neq k \Rightarrow A_j \cap A_k = \emptyset\}$. Note that given a pricing function $\pi_i : X_i \to \mathbb{R}$, the pricing based algorithm $\mathcal{A}$ assigns to agent i the set of items in $\arg\max_{x \in X_i} (\, v_i(x) - \pi(x)\,)$, which corresponds to the demand oracle in CAs. ∎

Prophet inequalities with matching constraint. We will apply our main result to the special case of prophet inequalities with a matching constraint for the edge-arrival and vertex arrival models. In the edge arrival model, edges in a graph appear online in a fixed but a priori unknown order, and each edge has a non-negative weight sampled from a known distribution. Upon the arrival of an edge, the algorithm decides irrevocably whether or not it wants to collect the edge. The collected edges should form a matching in the graph and the objective is to maximize the total weight of the collected set. In the vertex arrival model, the vertices of the graph appear online and, upon arrival, each vertex v_i reveals the edges (v_i, v_j) where $j < i$ (i.e., v_j has already arrived), along with their weights. The algorithm must decide whether to match v_i to some available v_j, collecting the weight of the edge (v_i, v_j), or to leave v_i unmatched. Note that an unmatched vertex v_i may later be matched by some v_j for $j > i$. As before, the goal is to maximize the total collected weight subject to the matching constraint.

For this problem, Ezra et al. [24] obtained algorithms with competitive ratios of $\frac{1}{2}$ and 0.337, for the vertex and edge arrival models, respectively. For the edge arrival model, this was subsequently improved to 0.344 by MacRury, Ma, and Grammel [38] using the same algorithm. Applying the above theorems, we obtain the following result which resolves an open problem of Ezra et al. [24].

Theorem 18. *For the prophet inequalities problem with matching constraints, there exists a pricing based algorithm with approximation $\frac{1}{2}$ in the vertex arrival model, and a pricing based algorithm with approximation* 0.344 *in the edge arrival model.*

Proof. Our result immediately applies to the edge arrival model with each agent corresponding to an edge. The set of outcomes is binary, i.e., we either accept the edge or do not. The valuation of agent i for the "accept" outcome is the weight of the corresponding edge and its valuation for the "reject" outcome is zero. The feasibility constraint $\mathcal{F}$ is that the set of agents assigned to the outcome "accept" should form a matching in the corresponding graph. The pricing algorithm corresponds to a threshold algorithm; for edge i, there is a threshold π_i such that denoting the edge's weight with w_i, we accept i if $w_i > \pi_i$ and reject if $w_i < \pi_i$. If $w_i = \pi_i$, we flip a coin with some pre-specified probability. [5]

For vertex arrival, the model studied by Ezra et al. [24] is *batched prophet inequalities*. Here, for each arriving vertex, the distribution of the edges corresponding to this vertex are sampled from some known distribution. While the distributions is not necessarily independent across the edges, the weights for each vertex are sampled independently of the previous vertices (i.e., there is independence across vertices). Our setting captures this problem as well, with each agent corresponding to a vertex. The set of outcomes for each vertex is the set of edges between this vertex and the previous vertices, as well as an additional "null" outcome of not accepting any vertex. □

6 Computational efficiency

In this section, we discuss how under standard assumptions, the reduction in Section 4 can be implemented efficiently, using access to a sampling oracle. We make two assumptions for the theorem. Firstly, we will assume that the distributions D_i are continuous. Secondly, we assume that the algorithm $\mathcal{A}^{\text{inp}}$ is *past-valuation independent*, that is, for each agent i, the assignment for i is based on v_i and $\mathcal{A}^{\text{inp}}_{<i}$, but does not (directly) depend on $v_1, \dots v_{i-1}$ (see Section 3 for a formal definition). Our main result is the following theorem, which is a more formal restatement of Theorem 2 from the introduction.

Theorem 19. *Let $\mathcal{A}^{\text{inp}}$ be a past-valuation independent algorithm for the prophet inequalities problem, and let $\mathbb{E}\left[\, v(\mathcal{A}^{\text{inp}}) \,\right]$ denote its expected social welfare. Let $\mathcal{F}$ denote a given combinatorial constraint satisfied by $\mathcal{A}^{\text{inp}}$. Assume further that the distributions D_i are continuous.* [6] *There exists an algorithm $\widehat{\mathcal{A}}^{\text{out}}$, that uses* $\text{poly}(n, \max_i |X_i|, \frac{1}{\epsilon})$ *samples, and can be implemented in* $\text{poly}(n, \max_i |X_i|, \frac{1}{\epsilon})$ *time with the following properties:*

- ***Approximately identical distribution.*** *The distribution of $\widehat{\mathcal{A}}^{\text{out}}$ approximately contains the distribution of $\mathcal{A}^{\text{inp}}$. Formally, letting $\mathcal{A}_i \in X_i$ denote the assignment made by algorithm $\mathcal{A}$ to agent i,*

$$\Pr\left[\, (\widehat{\mathcal{A}}^{\text{out}}_1, \dots, \widehat{\mathcal{A}}^{\text{out}}_n) = (x_1, \dots, x_n) \,\right] \geq (1-\epsilon)\Pr\left[\, (\mathcal{A}^{\text{inp}}_1, \dots, \mathcal{A}^{\text{inp}}_n) = (x_1, \dots, x_n) \,\right] \tag{6.8}$$

  for all $(x_1, \dots, x_n) \in \mathcal{X}$. [7]
- ***Feasibility.*** *$\widehat{\mathcal{A}}^{\text{out}} \in \mathcal{F}$ with probability 1.* [8]
- ***Approximately non-decreasing welfare.*** *The expected social welfare under algorithm $\widehat{\mathcal{A}}^{\text{out}}$ is at least $(1-\epsilon)$ times the expected social welfare under $\mathcal{A}^{\text{inp}}$. Formally, $\mathbb{E}\left[\, \mathbf{v}(\widehat{\mathcal{A}}^{\text{out}}) \,\right] \geq (1-\epsilon)\mathbb{E}\left[\, \mathbf{v}(\mathcal{A}^{\text{inp}}) \,\right]$. Moreover, for each agent i, $\mathbb{E}\left[\, v_i(\widehat{\mathcal{A}}^{\text{out}}_i) \,\right] \geq (1-\epsilon)\mathbb{E}\left[\, v_i(\mathcal{A}^{\text{inp}}_i) \,\right]$.*
- ***Pricing.*** *The algorithm $\widehat{\mathcal{A}}^{\text{out}}$ is pricing based.*

We note that the assumption on continuous distributions is standard and it can generally be lifted by adding a second "tie-breaking" coordinate to the valuation of each outcome, which is sampled from $[0, 1]$ (e.g., see [40]). The pricing function has a similar tie-breaking coordinate for each outcome. This ensures that the probability of a tie is zero as even if the utility of two outcomes is the same, their tie-breaking utility will be different with probability 1. Intuitively, this the same as adding a very small noise vector to each coordinate that makes the distributions continuous to avoid ties but does not have an effect if there are no ties.

As for the past-valuation independent assumption, we need this because as part of our algorithm, we need to estimate the probability distribution that the input algorithm would have given the currently assigned outcomes, i.e., $p_{\mathcal{A}^{\text{inp}}}(.; \widehat{\mathcal{A}}^{\text{out}}_{<i})$ as defined in Equation 4.5. Without the past-valuation assumption, estimating this would require us to repeatedly run the algorithm $\mathcal{A}^{\text{inp}}$ up to agent i using sampled valuations $v'_1, \dots, v'_i$, and restrict to those runs where $\mathcal{A}^{\text{inp}}_{<i}(\mathbf{v}') = \widehat{\mathcal{A}}^{\text{out}}_{<i}$. The number of samples required for this grows with $\frac{1}{\Pr\left[\, \mathcal{A}^{\text{inp}}_{<i}(\mathbf{v}') = \widehat{\mathcal{A}}^{\text{out}}_{<i} \,\right]}$ however, making the approach impractical if $\Pr\left[\, \mathcal{A}^{\text{inp}}_{<i}(\mathbf{v}') = \widehat{\mathcal{A}}^{\text{out}}_{<i} \,\right]$ is small. Note that we cannot simply ignore low-probability events because they can constitute a large fraction of the expected

[5] As mentioned in Section 3, in principle, our definition of a pricing algorithm allows for the probability of the coin to depend on the valuation function of v_i. For the algorithm obtained by our reduction however (see Lemma 10), the probability only depends on the set of tied outcomes. Since there are only two outcomes for edge arrival, this means that the tie-breaking probability is fixed.

[6] Specifically, we require that for any fixed $|X_i| - 1$ dimensional subspace, the probability that $v \sim D_i$ is in the subspace is zero.

[7] Note that the bound is asymmetric; $\Pr\left[\, \mathcal{A}^{\text{inp}}_{\leq n} = x_{\leq n} \,\right]$ is not necessarily $\geq (1-\epsilon)\Pr\left[\, \widehat{\mathcal{A}}^{\text{out}}_{\leq n} = x_{\leq n} \,\right]$.

[8] The constraint was not needed for Theorem 4 as it was guaranteed by the identical distribution condition. Since we can only prove approximately identical distribution here however, we state feasibility explicitly.

social welfare of $\mathcal{A}^{\text{inp}}$; even though the events happen with low probability, the expected social welfare when they happen may be very large. The past-valuation independent assumption resolves this issue by allowing us to simulate the algorithm $\mathcal{A}^{\text{inp}}$ without repeating the assignment procedure for agents 1 to $i-1$. While in principal an algorithm may not be past-valuation independent, most existing algorithms (including those discussed in Section 5) are past-valuation independent.

6.1 Algorithm Our algorithm proceeds in a similar way as Algorithm 1, with the difference that we solve the optimization problem (P1) using sampled data. As in Section 4.1, we focus on a single agent i and specify an assignment rule. As before, this means that the random variables $v_1, \ldots, v_{i-1}$ have been realized, and the algorithm has already assigned the outcomes $\mathcal{A}^{\text{out}}_{<i}$. As such, throughout the section, we will assume that $v_{<i}, \mathcal{A}^{\text{out}}_{<i}$ are fixed, deterministic values. As in Remark 5, when considering random variables for $v_{<i}, \mathcal{A}^{\text{out}}_{<i}$, the properties in this section will hold conditioned on $v_{<i}, \mathcal{A}^{\text{out}}_{<i}$.

Let m denote the number of samples we will use, which is a parameter to be specified later. For each agent i, we first estimate the probability vector $p_i(.) := p_{\mathcal{A}^{\text{inp}}}(., \widehat{\mathcal{A}}^{\text{out}}_{<i})$ (Definition 4.5), obtaining an estimate $\widehat{p}_i(.)$. Letting ϵ_i denote a parameter to be specified later, we will guarantee that $\left|\widehat{p}_i(x) - p_i(x)\right| \leq O(\epsilon_i)$, for all $x \in X_i$. In order to obtain this estimate, we run algorithm $\mathcal{A}^{\text{inp}}$ starting from agent i with $\mathcal{A}^{\text{inp}}_{<i}$ set to $\widehat{\mathcal{A}}^{\text{out}}_{<i}$ and with fake sample $v'_i \sim D_i$. Note that we do not need to rerun from the beginning because of the past-valuation independent assumption. Let $X_i^{\mathcal{F}} \subseteq X_i$ denote the set of feasible outcomes. Since we assumed that $\mathcal{A}^{\text{inp}}$ is feasible, $\widehat{p}_i(x) = 0$ for all $x \notin X_i^{\mathcal{F}}$.

Next, we add a small uniform noise to $\widehat{p}_i$ to obtain the distribution $\widetilde{p}_i$ defined as

$$\widetilde{p}_i(x) := (1-\epsilon_i)\widehat{p}_i(x) + \frac{\epsilon_i}{|X_i^{\mathcal{F}}|}\mathbb{1}\left\{x \in X_i^{\mathcal{F}}\right\}. \tag{6.9}$$

The reason behind this noise will become clear in the proofs. Intuitively however, the main reason is that in order to prove that $\widehat{\mathcal{A}}^{\text{out}}$ has high welfare, we will need to ensure a multiplicative bounds, as seen in the approximately identical distribution property of Theorem 19. This is because the expected welfare of the algorithm $\mathcal{A}^{\text{inp}}$ may be caused by a low probability event that a valuation function takes a very large value. Additive bounds are not suited to deal with this case as they may round down small probabilities to zero. The small uniform noise in (6.9) essentially solves this issue; if $p_i(x)$ is very small, the extra $\frac{\epsilon_i}{|X_i^{\mathcal{F}}|}$ probability will ensure that it is not ignored.

After obtaining $\widetilde{p}_i$, we take m samples from the distribution D_i, obtaining the data set $\widehat{D}_i$, and solve the optimization problem (P1) with $\widehat{D}_i$ and $\widetilde{p}_i$.[9] As discussed in Section 4.2, solving (P1) is equivalent to solving the optimization problem (LP1). We will solve the optimization problem over $X_i^{\mathcal{F}}$ only, as $p_i(x) = 0$ for $x \notin X_i^{\mathcal{F}}$. We let π_i denote the prices corresponding to the solution, which can be obtained as they are dual variables, and set $\pi_i(x) = +\infty$ for $x \notin X_i^{\mathcal{F}}$. After obtaining π_i, we observe the valuation function v_i, and choose the outcome $\arg\max_x (\, v_i(x) - \pi_i(x)\,)$. Ties can now be broken arbitrarily given the assumption that the distribution is continuous. A formal pseudocode of our approach is given in Algorithm 2.

Algorithm 2: Efficient reduction

Input: Input parameters $n, m, D_1, \ldots, D_n, \epsilon_1, \ldots, \epsilon_n, \mathcal{A}^{\text{inp}}$
Output: Assignment outcomes $\widehat{\mathcal{A}}^{\text{out}}_1, \ldots, \widehat{\mathcal{A}}^{\text{out}}_n$
for $i \in [n]$ **do**
  Set $\widehat{p}_i(x)$ for all $x \in X$ to be the estimate of $p_{\mathcal{A}^{\text{inp}}}(x; \widehat{\mathcal{A}}^{\text{out}}_{<i})$ obtained using m samples ;
  Set $\widetilde{p}_i(x)$ as in Equation (6.9) ;
  Obtain $\widehat{D}_i$ by sampling m times from the distribution D_i ;
  Solve the optimization problem (P1) with $(X_i^{\mathcal{F}}, \widehat{D}_i, \widetilde{p}_i)$, obtaining the solution $\widehat{A}^{\text{out}}_i$ with pricing function $\pi_i(x)$ for $x \notin X_i^{\mathcal{F}}$;
  Set $\pi_i(x) = +\infty$ for $x \in X_i^{\mathcal{F}}$;
  Observe the valuation function of the i-th agent v_i ;
  Assign the outcome $\widehat{\mathcal{A}}^{\text{out}}_i \in \arg\max_x (\, v_i(x) - \pi_i(x)\,)$, breaking ties arbitrarily ;
end

---

[9]We abuse notation and use $\widehat{D}_i$ to denote both the dataset, and the corresponding empirical distribution.

6.2 Proof of Theorem 19 To prove the theorem, we show that the assignment rules $\widehat{A}^{\text{out}}$ satisfy approximate versions of the properties of the assignment rule $\widehat{A}$ discussed in Section 4. Defining $\overline{A}_i^{\text{inp}}$ as $\overline{A}^{\text{inp}}_{\widehat{A}^{\text{out}}_{<i}}$ (Defintion (4.4)), we will show the following hold with high probability over the randomness of $\widehat{p}_i, \widehat{D}_i$.

1. **Approximately identical distribution.** $\Pr_{v\sim D_i}\left[\widehat{A}_i^{\text{out}}(v) = x_i\right] \geq (1-\epsilon)\Pr_{v\sim D_i}\left[\overline{A}_i^{\text{inp}}(v) = x_i\right]$ for all $x_i \in X_i$,
2. **Approximately non-decreasing welfare.** $\mathbb{E}_{v\sim D_i}\left[v(\widehat{A}_i^{\text{out}}(v))\right] \geq (1-\epsilon)\mathbb{E}_{v\sim D_i}\left[v(\overline{A}_i^{\text{inp}}(v))\right]$,
3. **Pricing.** A_i^{out} is pricing based, with a corresponding pricing function π_i.

We first show the approximately identical distribution for $\widehat{A}^{\text{out}}$ in Lemma 24. The main idea behind the proof is a uniform convergence argument based on bounding the VC-dimension of all pricing based assignment rules. Given a large enough (but polynomial) number of samples, with high probability, $\Pr_{v\sim\widehat{D}_i}[A(v) = x]$ is close to $\Pr_{v\sim D_i}[A(v) = x]$ for all pricing assignment rules A (Lemma 22 and Lemma 23). Since $\widehat{A}^{\text{out}}$ satisfies the constraint of the optimization problem (P1) with $p = \widetilde{p}_i$, and $\widetilde{p}_i$ is a good approximation of p_i (see Lemma 21), this proves that the approximately identical distribution condition for $\widehat{A}_i^{\text{out}}$. The approximately identical distribution property for the assignment rule $\widehat{A}^{\text{out}}$ leads to the corresponding property for $\widehat{\mathcal{A}}^{\text{out}}$, as required by Theorem 19 (Lemma 25).

We then use the approximately identical distribution property, together with the fact that $\widehat{A}_i^{\text{out}}$ is the solution to the optimization problem (P1) with $\widehat{D}_i$, to prove the approximately non-decreasing welfare property. Specifically, the approximately identical distribution property ensures that the probability distribution of $\widehat{A}_i^{\text{out}}$ over x is close to p_i. We show that by changing the output of the assignment rule $\overline{A}_i^{\text{inp}}$ with low probability, we can obtain a new "perturbed" assignment rule A' that has the same distribution over X_i as $\widehat{A}_i^{\text{out}}$ (Lemma 28). The perturbation approximately preserves the expected valuation of $\overline{A}_i^{\text{inp}}$ since it happens with low probability. Since $\widehat{A}_i^{\text{out}}$ is pricing based however, it is the optimal assignment rule for its distribution (Lemma 6), which means $\widehat{A}_i^{\text{out}}$ has a higher expected valuation than A'. Therefore, the expected valuation of $\widehat{A}_i^{\text{out}}$ is approximately larger than the expected valuation of $\overline{A}_i^{\text{inp}}$, implying that, with high probability, $\widehat{A}_i^{\text{out}}$ has the approximately non-decreasing welfare property (Lemma 29). Together with the approximately identical distribution property, this implies the approximately non-decreasing welfare property for $\widehat{\mathcal{A}}^{\text{out}}$, finishing the proof.

We proceed with a formal proof. We first discuss how to transform an additive ϵ error to a $(1-\epsilon)$ factor in the bounds. We begin with the following definition.

Definition 20. (ϵ-closeness) *Given two random variables A, B taking values in some set X, we say that A is ϵ-close to B if for all $x \in X$,*

$$\Pr[A = x] \geq (1-\epsilon)\Pr[B = x] .$$

Lemma 21. *Let A, B be random variables over a finite set X such that*

$$|\Pr[A = x] - \Pr[B = x]| \leq \frac{\epsilon}{|X|(1-\epsilon)}$$

for any $x \in X$. Define the random variable $\widetilde{A}$ as follows. With probability $1-\epsilon$, sample $\widetilde{A}$ from A. Otherwise, sample $\widetilde{A}$ uniformly at random from X. The random variable $\widetilde{A}$ is ϵ-close to B.

Proof. For any $x \in X$,

$$\begin{aligned}
\Pr\left[\widetilde{A} = x\right] &= (1-\epsilon)\Pr[A = x] + \frac{\epsilon}{|X|} && \textit{(Defintion of } \widetilde{A}\textit{)} \\
&\geq (1-\epsilon)\left(\Pr[B = x] - \frac{\epsilon}{|X|(1-\epsilon)}\right) + \frac{\epsilon}{|X|} && \textit{(Assumption on } A, B\textit{)} \\
&= (1-\epsilon)\Pr[B = x] .
\end{aligned}$$

□

We now bound the additive error caused by solving (LP1) with $\widehat{D}_i$ instead of D_i. Specifically, we show that with high probability over the randomness of $\widehat{D}_i$, all pricing based assignment rules have similar assignment distributions (over X_i) under D_i and $\widehat{D}_i$. We divide the proof into two parts. We first show this for a restricted class of pricing based assignment rules that break ties in lexicographical order (Lemma 22). We then generalize the argument to all pricing rules by bounding the effect of tie-breaking (Lemma 23).

Lemma 22. *Given a distribution D, Let $\widehat{D}$ be the empirical distribution formed by taking m independent samples from D. For any pricing function π, define A_π as the assignment rule that maps v to* $\arg\max_x (\, v(x) - \pi(x)\,)$*, breaking ties in lexicographical order. Let $\epsilon, \delta > 0$ be arbitrary parameters. There exists $M \le \mathrm{poly}(|X|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$ such that if $m \ge M$, with probability $1-\delta$ (over the randomness of $\widehat{D}$), for any $x \in X$ and any A_π* [10]

$$\left|\Pr_{v\sim\widehat{D}}\left[A_\pi(v) = x\right] - \Pr_{v\sim D}\left[A_\pi(v) = x\right]\right| \le \epsilon\,. \tag{6.10}$$

Proof. For any pricing vector π, and any $x, x' \in X$, let $f_{x,x',\pi}(v) := \mathbb{1}\{v(x) - \pi(x) \ge v(x') - \pi(x')\}$ denote the classifier that outputs 1 if and only if x is preferable to x' when using the price vector π, breaking ties in favor of x. We similarly define $f'_{x,x',\pi}(v) := \mathbb{1}\{v(x) - \pi(x) > v(x') - \pi(x')\}$, breaking ties in favor of x'. Define the set $F := \cup_{x,x',\pi}\{f_{x,x',\pi}, f'_{x,x',\pi}\}$. Let $F^\cap$ denote the set of all classifiers obtained by intersecting at most $|X|$ classifiers in F. In other words, $F^\cap$ is the set of classifiers f of the form $f(x) = f_1(x) \wedge \cdots \wedge f_k(x)$ for some $k \le |X|$.

Let d denote the VC-dimension of $F^\cap$. Using the fundamental theorem of statistical learning (Theorem 6.8 in [41]), as long as $m \ge O(\frac{d+\log(1/\delta)}{\varepsilon^2})$, with probability $1-\delta$ (over the randomness in sampling $\widehat{D}$), for any $f \in F^\cap$,

$$\left|\mathbb{E}_D\left[f\right] - \mathbb{E}_{\widehat{D}}\left[f\right]\right| \le \epsilon\,.$$

We note however that for any A_π and any x, the classifier $f(v) = \mathbb{1}\{A_\pi(v) = x\}$ is in $F^\cap$. This is because the event $\{\,A_\pi(v) = x\,\}$ is equivalent to x strictly being preferable to all x' that beat it lexicographically, and weakly preferable to x' that it beats lexicographically. Therefore, with probability $1-\delta$, for any $x \in X$ and any A_π,

$$\left|\Pr_{v\sim\widehat{D}}\left[A_\pi(v) = x\right] - \Pr_{v\sim D}\left[A_\pi(v) = x\right]\right| \le \epsilon\,.$$

It remains to bound the VC-dimension d. We first bound the VC-dimension of $F_{x,x'} := \cup_\pi\{f_{x,\pi,x'} \cup f'_{x,\pi,x'}\}$ for some fixed x, x'. Each $f \in F_{x,x'}$ is a half space in two dimensional space and therefore its VC-dimension is $O(1)$. It follows that the VC-dimension of F is at most $O(|X|^2\log(|X|))$ since it is the union of all sets $F_{x,x'}$ [5]. Letting d' denote the VC-dimension of F, the VC-dimension of $F^\cap$ is at most $O(|X|\log(|X|)d')$ because each classifier in $F^\cap$ is the intersection of at most $|X|$ classifiers in F [5]. ◻

Lemma 23. *Given a distribution D, let $\widehat{D}$ be the empirical distribution formed by taking m independent samples from D. Let $\epsilon, \delta > 0$ be arbitrary parameters. There exists $M \le \mathrm{poly}(|X|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$ such that if $m \ge M$, with probability $1-\delta$ (over the randomness of $\widehat{D}$), for any $x \in X$ and any pricing based assignment rule A*

$$\left|\Pr_{v\sim\widehat{D}}\left[A(v) = x\right] - \Pr_{v\sim D}\left[A(v) = x\right]\right| \le \epsilon\,. \tag{6.11}$$

Proof. Let π be the corresponding pricing vector for A and consider the assignment rule A_π. We know from Lemma 22 that Equation (6.11) holds with A_π instead of A. The only difference between the two assignment rules is how they break ties when two outcomes have the same utility. Tie-breaking does not affect $\Pr_{v\sim D}\left[A(v) = x\right]$ however as the distribution is continuous. We will show that tie-breaking does not have a large effect on $\Pr_{v\sim\widehat{D}}\left[A(v) = x\right]$ either.

Formally, consider any outcomes x and x'. Let $\mathrm{tie}_{x,x'}(v)$ denote the indicator random variable that x and x' are tied for the valuation function v and the pricing function π, i.e., $\mathrm{tie}_{x,x'}(v) = \mathbb{1}\{v(x) - \pi(x) = v(x') - \pi(x')\}$. It is clear that $A_\pi(v)$ and $A(v)$ can only be different if there is a tie for two classes. Therefore,

$$|\,\mathbb{1}\{A_\pi(v) = x\} - \mathbb{1}\{A(v) = x\}\,| \le \sum_{x'\ne x}\mathbb{1}\{\mathrm{tie}_{x,x'}(v)\}\,.$$

It follows that

$$\begin{aligned}\left|\mathbb{E}_{v\sim\widehat{D}}\left[\mathbb{1}\{A_\pi(v) = x\}\right] - \mathbb{E}_{v\sim\widehat{D}}\left[\mathbb{1}\{A(v) = x\}\right]\right| &\le \mathbb{E}_{v\sim\widehat{D}}\left[|\,\mathbb{1}\{A_\pi(v) = x\} - \mathbb{1}\{A(v) = x\}\,|\right]\\ &\le \sum_{x'\ne x}\mathbb{E}_{v\sim\widehat{D}}\left[\mathbb{1}\{\mathrm{tie}_{x,x'}(v)\}\right]\,.\end{aligned}$$

Note however that if $\mathrm{tie}_{x,x'}(v) = 1$, then

$$v(x) - v(x') = \pi(x) - \pi(x')\,.$$

[10] The dataset $\widehat{D}$ in Equation (6.10) is fixed and the randomness is over the random draw of v, and the internal randomness of A_π, but not the randomness in $\widehat{D}$. In other words, defining the set of "good samples" $\widehat{D}$ as those satisfying (6.10), the lemma states that $\widehat{D}$ will be "good" with probability at least $1-\delta$.

Therefore, if for any x', there are two valuation functions $v, v' \in \widehat{D}$ such that $\text{tie}_{x,x'}(v) = 1$ and $\text{tie}_{x,x'}(v') = 1$, then

$$v(x) - v(x') = v'(x) - v'(x') .$$

The probability that this happens is zero however because D was assumed to be continuous and $\widehat{D}$ was formed by taking samples from D. More specifically, assuming we sample the valuation functions in $\widehat{D}$ one by one and sample v' before v, then after sampling v' the subspace $\{v : v(x) - v(x') = v'(x) - v'(x')\}$ is fixed and the probability that we sample from it is zero because it is $|X| - 1$ dimensional.

It follows that

$$\sum_{x' \neq x} \mathbb{E}_{v \sim \widehat{D}} \left[\mathbb{1}\left\{\text{tie}_{x,x'}(v)\right\}\right] \leq \sum_{x'} \frac{1}{m} = \frac{|X|}{m} .$$

Therefore, as long as $m \geq \frac{2|X|}{\epsilon}$, for any x, x' and A_π.

$$\left|\Pr_{v \sim \widehat{D}}\left[A_\pi(v) = x\right] - \Pr_{v \sim \widehat{D}}\left[A(v) = x\right]\right| = \left|\mathbb{E}_{v \sim \widehat{D}}\left[\mathbb{1}\left\{A_\pi(v) = x\right\}\right] - \mathbb{E}_{v \sim \widehat{D}}\left[\mathbb{1}\left\{A(v) = x\right\}\right]\right| \leq \frac{\epsilon}{2} .$$

Note however that by Lemma 22, as long as $m \geq M$ for some $M \leq \text{poly}(|X|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$, with probability $1 - \delta$,

$$\left|\Pr_{v \sim D}\left[A_\pi(v) = x\right] - \Pr_{v \sim \widehat{D}}\left[A_\pi(v) = x\right]\right| \leq \epsilon/2 .$$

Since $\Pr_{v \sim D}\left[A_\pi(v) = x\right] = \Pr_{v \sim D}\left[A(v) = x\right]$, the claim follows. ∎

Using the above results, we obtain the following important lemma, which shows that with high-probability, the assignment rule $\widehat{A}_i^{\text{out}}$ has the approximately identical distribution property.

Lemma 24. *Let $i \in [n]$ and assume that $\widehat{\mathcal{A}}_{<i}^{\text{out}}$ is a fixed, deterministic value. Let $\epsilon > 0$ be an arbitrary parameter and set $\epsilon_i = \epsilon/2$. There exists $M \leq \text{poly}(|X_i|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$ such that when $m \geq M$, with probability $1 - \delta$ (over the randomness in calculation of $\widehat{A}_i^{\text{out}}$), for all $x \in X_i$,* [11]

$$\Pr_{v_i \sim D_i}\left[\widehat{A}_i^{\text{out}}(v_i) = x\right] \geq (1 - \epsilon) p_i(x) . \tag{6.12}$$

Proof. We assume without loss of generality that $X_i = X_i^{\mathcal{F}}$ because for $x \notin X_i^{\mathcal{F}}$, both sides of the inequality are zero. Set M to be large enough such that with probability $1 - \delta/2$, for all x we have

$$\left|\widehat{p}_i(x) - p_i(x)\right| \leq \frac{\epsilon_i}{|X_i|}.$$

A standard Chernoff argument implies that this can be done using $M \leq \text{poly}(|X|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$. Given Lemma 21, with probability $1 - \delta/2$, for any x,

$$\widetilde{p}_i(x) \geq (1 - \epsilon_i) p_i(x). \tag{6.13}$$

Assume further that M is large enough such that with probability $1 - \delta/2$, for all pricing based assignment rules A, we have

$$\left|\Pr_{v_i \sim \widehat{D}_i}\left[A(v_i) = x\right] - \Pr_{v_i \sim D_i}\left[A(v_i) = x\right]\right| \leq \epsilon_i^2/|X|.$$

By Lemma 23, this can be done using $M \leq \text{poly}(|X_i|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$. Since $\widehat{A}_i^{\text{out}}$ is feasible for (P1) with (D, p) set to $(\widehat{D}_i, \widetilde{p}_i)$,

$$\Pr_{v_i \sim \widehat{D}_i}\left[\widehat{A}_i^{\text{out}}(v_i) = x\right] = \widetilde{p}_i(x).$$

Therefore, with probability $1 - \delta$,

$$\Pr_{v_i \sim D_i}\left[\widehat{A}_i^{\text{out}}(v_i) = x\right] \geq \widetilde{p}_i(x) - \epsilon_i^2/|X| \geq (1 - \epsilon_i)\widetilde{p}_i(x),$$

[11] In Equation (6.12), $\widehat{A}_i^{\text{out}}$ is fixed and the probabilities are over the internal randomness of $\widehat{A}_i^{\text{out}}$ and v_i, but not over the randomness of calculating $\widehat{A}_i^{\text{out}}$ (i.e., the randomness of $\widehat{p}_i, \widehat{D}_i$). Stated differently, defining any fixed assignment rule $\widehat{A}_i^{\text{out}}$ "good" if it satisfies (6.12), the lemma states that the $\widehat{A}_i^{\text{out}}$ obtained by our algorithm is "good" with high probability.

where the second inequality follows from the fact that $\widetilde{p}_i(x) \geq \epsilon_i/|X|$ for all x. Therefore, for some $M \leq \mathrm{poly}(|X_i|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$, with probability $(1-\delta)$,

$$\Pr_{v_i \sim D_i}\left[\widehat{A}_i^{\mathrm{out}}(v_i) = x\right] \geq (1-\epsilon_i)\widetilde{p}_i(x) \geq (1-\epsilon_i)^2 p_i(x) \geq (1-2\epsilon_i)p_i(x) = (1-\epsilon)p_i(x)$$

where the second inequality follows from (6.13). Therefore, we have obtained Equation (6.12), finishing the proof. ∎

Using the above lemma, we can prove the approximately identical distribution property for $\widehat{\mathcal{A}}^{\mathrm{out}}$.

Lemma 25. *Let $\epsilon > 0$ be an arbitrary parameter. There exists $M \leq \mathrm{poly}(n, \max_i |X_i|, \frac{1}{\epsilon})$ and $\epsilon_1, \ldots, \epsilon_n$ such that if $m \geq M$, for any i and $(x_1, \ldots, x_i) \in X_1 \times \cdots \times X_i$,*

$$\Pr\left[\widehat{\mathcal{A}}_{\leq i}^{\mathrm{out}} = x_{\leq i}\right] \geq (1-\epsilon)\Pr\left[\mathcal{A}_{\leq i}^{\mathrm{inp}} = x_{\leq i}\right].$$

Proof. We start with the following claim.

Claim 26. *For any i, there exists $M \leq \mathrm{poly}(|X_i|, \frac{1}{\epsilon})$ such that when $m \geq M$, for any $x_{<i}, x$,*

$$\Pr\left[\widehat{\mathcal{A}}_i^{\mathrm{out}} = x \mid \widehat{\mathcal{A}}_{<i}^{\mathrm{out}} = x_{<i}\right] \geq (1-\epsilon)\Pr\left[\mathcal{A}_i^{\mathrm{inp}} = x \mid \mathcal{A}_{<i}^{\mathrm{inp}} = x_{<i}\right]$$

Proof. Set $\delta = \epsilon$. By Lemma 24, there exists $M \leq \mathrm{poly}(|X_i|, \frac{1}{\epsilon})$ such that for any i, with probability $1-\delta$ (over the randomness in calculating $\widehat{A}_i^{\mathrm{out}}$) we have

$$\Pr_{v_1,\ldots,v_i}\left[\widehat{A}_i^{\mathrm{out}}(v_i) = x \mid A_{<i}^{\mathrm{out}} = x_{<i}\right] \geq (1-\epsilon)p_{\mathcal{A}^{\mathrm{inp}}}(x; x_{<i}). \tag{6.14}$$

It follows that

$$\Pr\left[\widehat{\mathcal{A}}_i^{\mathrm{out}} = x \mid \widehat{\mathcal{A}}_{<i}^{\mathrm{out}} = x_{<i}\right] \geq (1-\delta)(1-\epsilon)p_{\mathcal{A}^{\mathrm{inp}}}(x; x_{<i}) \geq (1-2\epsilon)p_{\mathcal{A}^{\mathrm{inp}}}(x; x_{<i}).$$

Since $p_{\mathcal{A}^{\mathrm{inp}}}(x; x_{<i}) = \Pr\left[\mathcal{A}_i^{\mathrm{inp}} = x \mid \mathcal{A}_{<i}^{\mathrm{inp}} = x_{<i}\right]$, replacing ϵ with $\epsilon/2$ proves the claim. ∎

Given this claim, there exists $M \leq \mathrm{poly}(n, \max_i |X_i|, \frac{1}{\epsilon})$ such that if $m \geq M$, for any $i, x_{<i}, x_i$,

$$\Pr\left[\widehat{\mathcal{A}}_i^{\mathrm{out}} = x_i \mid \widehat{\mathcal{A}}_{<i}^{\mathrm{out}} = x_{<i}\right] \geq (1-\epsilon/n)\Pr\left[\mathcal{A}_i^{\mathrm{inp}} = x_i \mid \mathcal{A}_{<i}^{\mathrm{inp}} = x_{<i}\right]. \tag{6.15}$$

We prove by induction on i that

$$\Pr\left[\widehat{\mathcal{A}}_{\leq i}^{\mathrm{out}} = x_{\leq i}\right] \geq (1-i\epsilon/n)\Pr\left[\mathcal{A}_{\leq i}^{\mathrm{inp}} = x_{\leq i}\right].$$

The base case of $i = 0$ holds trivially. Assuming the claim holds for $i-1$,

$$\begin{aligned}
\Pr\left[\widehat{\mathcal{A}}_{\leq i}^{\mathrm{out}} = x_{\leq i}\right] &= \Pr\left[\widehat{\mathcal{A}}_{<i}^{\mathrm{out}} = x_{<i}\right]\Pr\left[\widehat{\mathcal{A}}_i^{\mathrm{out}} = x_i \mid \widehat{\mathcal{A}}_{<i}^{\mathrm{out}} = x_{<i}\right] \\
&\geq (1-\frac{(i-1)\epsilon}{n})\Pr\left[\mathcal{A}_{<i}^{\mathrm{inp}} = x_{<i}\right]\Pr\left[\widehat{\mathcal{A}}_i^{\mathrm{out}} = x_i \mid \widehat{\mathcal{A}}_{<i}^{\mathrm{out}} = x_{<i}\right] && \textit{(Induction hypothesis)} \\
&\geq (1-\frac{(i-1)\epsilon}{n})(1-\frac{\epsilon}{n})\Pr\left[\mathcal{A}_{<i}^{\mathrm{inp}} = x_{<i}\right]\Pr\left[\mathcal{A}_i^{\mathrm{inp}} = x_i \mid \mathcal{A}_{<i}^{\mathrm{inp}} = x_{<i}\right] && \textit{(Equation (6.15))} \\
&\geq (1-\frac{i\epsilon}{n})\Pr\left[\mathcal{A}_{\leq i}^{\mathrm{inp}} = x_{\leq i}\right],
\end{aligned}$$

finishing the proof. ∎

Finally, we will focus on proving that the approximately non-decreasing welfare property. To do this, we first provide an alternative view of ϵ-closeness. Given random variables A, B where A is ϵ-close to B, we show that we can obtain the random variable A by first sampling from B, and then changing the obtained sample with some small probability. Formally, we define the notion of a *low-probability switch* as follows.

Definition 27. (Low-probability switch) *Given a distribution q over X, a value $x \in X$, and a parameter $p \in [0, 1]$, we define the low probability switch of x with q, p, denoted by $\mathrm{LPS}(x, q, p)$ as follows. Sample the random variable $Y \sim \mathrm{Bernoulli}(p)$. If $Y = 0$, set $\mathrm{LPS}(x, q, p) := A$. Otherwise, sample $\mathrm{LPS}(x, q, p)$ from q.*

Lemma 28. *Let A, B be two random variables taking values in a set X and assume that B is ϵ-close to A. There exists a distribution q such that A has the same distribution as $\mathrm{LPS}(B, q, \epsilon)$.*

Proof. Define $q(x)$ as

$$q(x) = \frac{1}{\epsilon}\left(\Pr\left[A = x\right] - (1-\epsilon)\Pr\left[B = x\right]\right).$$

Note that $q(x) \geq 0$ by the assumption on A and B and that

$$\sum_x q(x) = \frac{1}{\epsilon}\left(\sum_x \Pr\left[A = x\right] - (1-\epsilon)\sum_x \Pr\left[B = x\right]\right) = 1.$$

Therefore, q is a probability distribution over X. By definition of q,

$$\Pr\left[\mathrm{LPS}(B, q, \epsilon) = x\right] = (1-\epsilon)\Pr\left[B = x\right] + \epsilon q(x) = \Pr\left[A = x\right]. \qquad \textit{(Definition of q)}$$

□

Lemma 29. *Let $i \in [n]$ and condition on the event $\widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}$ for some $x_{<i} \in X_1 \times \cdots \times X_{i-1}$. There exists $M \leq \mathrm{poly}(|X_i|, \frac{1}{\epsilon}, \log(\frac{1}{\delta}))$ such that when $m \geq M$, with probability $1-\delta$ (over the randomness in calculation of $\widehat{A}^{\text{out}}_i$),*

$$\mathbb{E}_{v\sim D_i}\left[v(\widehat{A}^{\text{out}}_i(v))\right] \geq (1-\epsilon)\mathbb{E}_{v\sim D_i}\left[v(\overline{A}^{\text{inp}}_{x_{<i}}(v))\right].$$

Proof. By Lemma 24, conditioned on $\widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}$, with probability $1-\delta$ (over the randomness of calculating $\widehat{A}^{\text{out}}_i$),

$$\Pr_{v\sim D_i}\left[\widehat{A}^{\text{out}}(v) = x\right] \geq (1-\epsilon) p_{\mathcal{A}^{\text{inp}}}(x; x_{<i}),$$

By Lemma 12 however,

$$\Pr_{v_i\sim D_i}\left[\overline{A}^{\text{inp}}_{x_{<i}}(v_i) = x\right] = p_{\mathcal{A}^{\text{inp}}}(x; x_{<i}).$$

Therefore, by Lemma 28, conditioned on $\widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}$, with probability $1-\delta$ (over the randomness of calculating $\widehat{A}^{\text{out}}_i$), there exists a distribution q_i such that

$$\Pr_{v\sim D_i}\left[\widehat{A}^{\text{out}}(v) = x\right] = \Pr_{v\sim D_i}\left[\mathrm{LPS}(\overline{A}^{\text{inp}}(v), q_i, \epsilon)\right].$$

Define the assignment rule A' as $A'(v) := \mathrm{LPS}(\overline{A}^{\text{inp}}_{x_{<i}}(v), q_i, \epsilon)$. On one hand, since the valuation functions are always non-negative,

$$\mathbb{E}_{v\sim D_i}\left[v(A'(v))\right] \geq (1-\epsilon)\mathbb{E}_{v\sim D_i}\left[v(\overline{A}^{\text{inp}}_{x_{<i}}(v))\right].$$

On the other hand, A'_i is feasible for the optimization problem (P1) with $D = D_i$ and p set to $p(x) = \Pr_{v\sim D_i}\left[\widehat{A}^{\text{out}}(v) = x\right]$. Since $\widehat{A}_i$ is also feasible for this problem, and it is pricing based, Lemma 6 implies that

$$\mathbb{E}_{v\sim D_i}\left[v(\widehat{A}(v))\right] \geq \mathbb{E}_{v\sim D_i}\left[v(A'(v))\right].$$

Therefore, conditioned on $\widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}$, with probability $1-\delta$ (over the randomness of calculating $\widehat{A}^{\text{out}}_i$),

$$\mathbb{E}_{v\sim D_i}\left[v(\widehat{A}^{\text{out}}_i(v))\right] \geq (1-\epsilon)\mathbb{E}_{v\sim D_i}\left[v(\overline{A}^{\text{inp}}_{x_{<i}}(v))\right],$$

as claimed. □

We now prove the approximately non-decreasing welfare property.

Lemma 30. *For any $\epsilon > 0$, there exists $M \leq \mathrm{poly}(n, \max_i |X_i|, \frac{1}{\epsilon})$ such that if $m \geq M$, such that for any i,*

$$\mathbb{E}\left[\, v_i(\widehat{\mathcal{A}}^{\text{out}}_i)\,\right] \geq (1-\epsilon)\mathbb{E}\left[\, v_i(\mathcal{A}^{\text{inp}}_i)\,\right].$$

Proof. By Lemma 25 there exists $M \le \text{poly}(n, \max_i |X_i|, \frac{1}{\epsilon})$ such that if $m \ge M$, for any $x_{<i}$,

$$\Pr\left[\widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}\right] \ge (1-\epsilon)\Pr\left[\mathcal{A}^{\text{inp}}_{<i} = x_{<i}\right]. \tag{6.16}$$

We will show that for any $x_{<i}$,

$$\mathbb{E}\left[v_i(\widehat{\mathcal{A}}^{\text{out}}_i) \mid \widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}\right] \ge (1-\epsilon)\mathbb{E}\left[v_i(\mathcal{A}^{\text{inp}}_i) \mid \mathcal{A}^{\text{inp}}_{<i} = x_{<i}\right]. \tag{6.17}$$

Assuming this is proved, then it follows that

$$\begin{aligned}
\mathbb{E}\left[v_i(\widehat{\mathcal{A}}^{\text{out}}_i)\right] &= \sum_{x_{<i}} \Pr\left[\widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}\right]\mathbb{E}\left[v_i(\widehat{\mathcal{A}}^{\text{out}}_i) \mid \widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}\right] \\
&\ge \sum_{x_{<i}} (1-\epsilon)^2 \Pr\left[\mathcal{A}^{\text{inp}}_{<i} = x_{<i}\right]\mathbb{E}\left[v_i(\mathcal{A}^{\text{inp}}_i) \mid \mathcal{A}^{\text{inp}}_{<i} = x_{<i}\right] && \textit{(Equations (6.16) and (6.17))} \\
&\ge (1-2\epsilon)\mathbb{E}\left[v_i(\mathcal{A}^{\text{inp}}_i)\right].
\end{aligned}$$

Replacing ϵ with $\epsilon/2$ finishes the proof.

It remains to prove (6.17). By Lemma 15,

$$\mathbb{E}\left[v_i(\mathcal{A}^{\text{inp}}_i) \mid \mathcal{A}^{\text{inp}}_{<i} = x_{<i}\right] = \mathbb{E}_{v_i \sim D_i}\left[v_i(\overline{A}^{\text{inp}}_{x_{<i}}(v_i))\right], \tag{6.18}$$

By definition of $\widehat{\mathcal{A}}^{\text{out}}$,

$$\mathbb{E}\left[v_i(\widehat{\mathcal{A}}^{\text{out}}_i) \mid \widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}\right] = \mathbb{E}_{\widehat{p}_i, \widehat{D}_i, v_i}\left[v_i(\widehat{A}^{\text{out}}(v_i)) \mid \widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}\right], \tag{6.19}$$

Note that in the right-hand side, the assignment rule $\widehat{A}^{\text{out}}_i$ is itself random as it is the solution to the (P1) with the (random) parameters $\widehat{p}_i, \widehat{D}_i$. By Lemma 29 however, with probability $1-\delta$ over the randomness of $\widehat{p}_i, \widehat{D}_i$,

$$\mathbb{E}_{v \sim D_i}\left[v(\widehat{A}^{\text{out}}_i(v))\right] \ge (1-\epsilon/2)\mathbb{E}_{v \sim D_i}\left[v(\overline{A}^{\text{inp}}_{x_{<i}}(v))\right],$$

Therefore, setting $\delta = \epsilon/2$,

$$\mathbb{E}_{\widehat{p}_i, \widehat{D}_i, v_i}\left[v_i(\widehat{A}^{\text{out}}(v_i)) \mid \widehat{\mathcal{A}}^{\text{out}}_{<i} = x_{<i}\right] \ge (1-\delta)(1-\epsilon/2)\mathbb{E}_{v \sim D_i}\left[v(\overline{A}^{\text{inp}}_{x_{<i}}(v))\right] \ge (1-\epsilon)\mathbb{E}_{v \sim D_i}\left[v(\overline{A}^{\text{inp}}_{x_{<i}}(v))\right]$$

Combined with Equations (6.18) and (6.19), this proves Equation (6.17), finishing the proof. □

Proof. [Proof of Theorem 19] We use Algorithm 2. The approximately identical distribution property follows from Lemma 24. The approximately non-decreasing welfare property follows from Lemma 29. The pricing based property follows from Lemma 6, and the feasibility follows from the fact that the algorithm sets the threshold $+\infty$ for infeasible outcomes. □

7 Conclusions

In this paper, we introduce a black-box reduction that converts any algorithm for prophet inequalities into a pricing-based algorithm, solving an important open problem in the literature. The reduction maintains the distribution of the assigned outcomes and ensures that the algorithm achieves at least the same expected social welfare.

Going forward, there are a number of interesting directions for future work. Firstly, for the specific case of combinatorial auctions, our polynomial-time computation depends on the number of different outcomes which is exponential in the number of items. For the general case of bundle prices, this is an inherent limitation for any algorithm that explicitly calculates all prices. It would be interesting to resolve this issue however by either considering per-item prices or using prices implicitly. Additionally, it would be interesting to extend our results, possibly with extra assumptions or a loss in approximation, to prices that are more structured and have nice properties such as being anonymous.

8 Acknowledgements

We thank the anonymous reviewers for their valuable feedback. The work is partially supported by DARPA QuICC, NSF AF:Small #2218678, and NSF AF:Small #2114269, and Royal Society grant IES\R2\222170.

References

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

[2] S. Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. *SIAM Journal on Computing*, 43(2):930–972, 2014.

[3] P. D. Azar, R. Kleinberg, and S. M. Weinberg. Prophet inequalities with limited information. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 1358–1377. SIAM, 2014.

[4] C. Bechtel, S. Dughmi, and N. Patel. Delegated pandora's box. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 666–693, 2022.

[5] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Learnability and the vapnik-chervonenkis dimension. *Journal of the ACM (JACM)*, 36(4):929–965, 1989.

[6] L. Blumrosen and N. Nisan. Combinatorial auctions. *Chapter 11 in the book "Algorithmic Game Theory". Editors: N. Nisan, T. Roughgarden, E. Tardos and V. Vazirani.*, pages 267–300, 2007.

[7] Y. Cai, C. Daskalakis, and S. M. Weinberg. Understanding incentives: Mechanism design becomes algorithm design. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 618–627. IEEE, 2013.

[8] C. Caramanis, P. Dütting, M. Faw, F. Fusco, P. Lazos, S. Leonardi, O. Papadigenopoulos, E. Pountourakis, and R. Reiffenhäuser. Single-sample prophet inequalities via greedy-ordered selection. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1298–1325. SIAM, 2022.

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

[10] S. Chawla, N. Immorlica, and B. Lucier. On the limits of black-box reductions in mechanism design. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 435–448, 2012.

[11] P.-A. Chiappori, R. J. McCann, and L. P. Nesheim. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Economic Theory*, pages 317–354, 2010.

[12] J. Correa, A. Cristi, B. Epstein, and J. A. Soto. Sample-driven optimal stopping: From the secretary problem to the iid prophet inequality. *Mathematics of Operations Research*, 2023.

[13] J. Correa, P. Dütting, F. Fischer, and K. Schewior. Prophet inequalities for iid random variables from an unknown distribution. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 3–17, 2019.

[14] J. Correa, P. Foncea, D. Pizarro, and V. Verdugo. From pricing to prophets, and back! *Operations Research Letters*, 47(1):25–29, 2019.

[15] J. Correa, R. Saona, and B. Ziliotto. Prophet secretary through blind strategies. *Mathematical Programming*, 190(1-2):483–521, 2021.

[16] J. R. Correa and A. Cristi. A constant factor prophet inequality for online combinatorial auctions. In B. Saha and R. 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.

[17] S. Dughmi and T. Roughgarden. Black-box randomized reductions in algorithmic mechanism design. *SIAM Journal on Computing*, 43(1):312–336, 2014.

[18] P. Dütting, M. Feldman, T. Kesselheim, and B. Lucier. Prophet inequalities made easy: Stochastic optimization by pricing non-stochastic inputs. In C. Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 540–551. IEEE Computer Society, 2017.

[19] P. Dütting, M. Feldman, T. Kesselheim, and B. Lucier. Prophet inequalities made easy: Stochastic optimization by pricing nonstochastic inputs. *SIAM Journal on Computing*, 49(3):540–582, 2020.

[20] P. Dütting, T. Kesselheim, and B. Lucier. An O(log log m) prophet inequality for subadditive combinatorial auctions. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020, Durham, NC, USA, November 16-19, 2020*, pages 306–317. IEEE, 2020.

[21] S. Ehsani, M. Hajiaghayi, T. Kesselheim, and S. Singla. Prophet secretary for combinatorial auctions and matroids. In *Proceedings of the twenty-ninth annual acm-siam symposium on discrete algorithms*, pages 700–714. SIAM, 2018.

[22] H. Esfandiari, M. Hajiaghayi, V. Liaghat, and M. Monemizadeh. Prophet secretary. *SIAM Journal on Discrete Mathematics*, 31(3):1685–1701, 2017.

[23] T. Ezra, M. Feldman, N. Gravin, and Z. G. 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.

[24] T. Ezra, M. Feldman, N. Gravin, and Z. G. Tang. Prophet matching with general arrivals. *Mathematics of Operations Research*, 47(2):878–898, 2022.

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

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

[27] N. Gravin and H. 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.

[28] N. E. Gretsky, J. M. Ostroy, and W. R. Zame. The nonatomic assignment model. *Economic Theory*, 2:103–127, 1992.

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

[30] M. T. Hajiaghayi, R. Kleinberg, and T. Sandholm. Automated online mechanism design and prophet inequalities. In *AAAI*, volume 7, pages 58–65, 2007.

[31] M. T. Hajiaghayi, R. D. Kleinberg, and T. Sandholm. Automated online mechanism design and prophet inequalities. In *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, July 22-26, 2007, Vancouver, British Columbia, Canada*, pages 58–65. AAAI Press, 2007.

[32] J. D. Hartline, R. Kleinberg, and A. Malekian. Bayesian incentive compatibility via matchings. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 734–747. SIAM, 2011.

[33] J. D. Hartline and B. Lucier. Bayesian algorithmic mechanism design. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 301–310, 2010.

[34] J. Kleinberg and R. Kleinberg. Delegated search approximates efficient search. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 287–302, 2018.

[35] R. Kleinberg and S. M. Weinberg. Matroid prophet inequalities. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 123–136, 2012.

[36] U. Krengel and L. Sucheston. Semiamarts and finite values. *Bull. Amer. Math. Soc.*, 83:745–747, 1977.

[37] U. Krengel and L. Sucheston. On semiamarts, amarts, and processes with finite value. *Adv. in Probability 4:197–266*, 4:197–266, 1978.

[38] C. MacRury, W. Ma, and N. Grammel. On (random-order) online contention resolution schemes for the matching polytope of (bipartite) graphs. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1995–2014. SIAM, 2023.

[39] A. Rubinstein. Beyond matroids: Secretary problem and prophet inequality with general constraints. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 324–332, 2016.

[40] A. Rubinstein, J. Z. Wang, and S. M. Weinberg. Optimal single-choice prophet inequalities from samples. *arXiv preprint arXiv:1911.07945*, 2019.

[41] S. Shalev-Shwartz and S. Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

[42] L. S. Shapley and M. Shubik. The assignment game i: The core. *International Journal of game theory*, 1:111–130, 1971.

[43] C. Villani. *Optimal transport: old and new*, volume 338. Springer, 2009.

[44] Q. Yan. Mechanism design via correlation gap. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 710–719. SIAM, 2011.

A Extenstion to prophet secretary

In the prophet secretary problem, the arrival order of the agents is not fixed beforehand and instead is chosen uniformly at random. Letting $\sigma : [n] \to [n]$ denote this distribution, this is equivalent to assuming that the input is $D_{\sigma(1)}, \ldots, D_{\sigma(n)}$ instead of $D_1, \ldots, D_n$. Since our results hold for any fixed realization of σ, they hold in expectation over σ as well. For instance, by Theorem 1, for any $\mathbf{x} \in \mathcal{X}$[12],

$$\Pr\left[\mathcal{A}^{\text{inp}} = \mathbf{x} \mid \sigma\right] = \Pr\left[\mathcal{A}^{\text{out}} = \mathbf{x} \mid \sigma\right],$$

which by iterated expectation implies

$$\Pr\left[\mathcal{A}^{\text{inp}} = \mathbf{x}\right] = \Pr\left[\mathcal{A}^{\text{out}} = \mathbf{x}\right].$$

Simliarly, $\mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{out}}) \mid \sigma\right] \geq \mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{inp}}) \mid \sigma\right]$, which implies $\mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{out}})\right] \geq \mathbb{E}\left[\mathbf{v}(\mathcal{A}^{\text{inp}})\right]$ by iterated expectation. The same argument applies for the results in Theorem 2.

B Omitted proofs

B.1 Proof of Lemma 8

Proof. It is clear that for any assignment rule A, the tuple $(v, A(v))$ where $v \sim D$ has a joint distribution over $V \times X$ with marginal D on V and $A(D)$ on X. Letting $\gamma_{A,D}$ denote this distribution, $\mathbb{E}_{(v,x)\sim\gamma_{A,D}}[v(x)] = \mathbb{E}_{v\sim D}[v(A(v))]$ by definition of $\gamma_{A,D}$. □

B.2 Proof of Lemma 10

Proof. Assume that A is a pricing based assignment rule. Define $\psi(v) := \max_x (v(x) - \pi(x))$. By definition, for any x,

$$\psi(v) + \pi(x) \geq v(x) - \pi(x) + \pi(x) = v(x).$$

Additionally, For any v we know that $A(v)$ is going to be in $\arg\max_{x'} (v(x') - \pi(x'))$, which implies that $v(x) - \pi(x) = \psi(v)$ with probability 1 when $(v, x) \sim \gamma_A$.

Conversely, assume that γ is a pricing based coupling. Let π be the corresponding pricing function. We will provide an assignment rule A with the pricing function π. If there are no ties, A will simply map v to $\arg\max_x (v(x) - \pi(x))$. In case of ties, we will replicate the same tie-breaking distribution as γ.

Formally, for any v, define $\text{tieset}_v \subseteq X$ as $\text{tieset}_v := \arg\max_{x'} (v(x) - \pi(x))$. Note that since π is fixed, tieset_v is a function of v. For any $X' \subseteq X$, define the probability distribution $q_{\gamma,X'}(.)$ over X as the tie-breaking distribution of γ when $\text{tieset}_v = X'$, i.e.,

$$q_{\gamma,X'}(\tilde{x}) = \Pr_{(v,x)\sim\gamma}\left[x = \tilde{x} \mid X' = \text{tieset}_v\right].$$

We note that in the above definition, $x' \in X$ is a fixed value. Given any v, we sample $A(v)$ from the distribution $q_{\gamma,\text{tieset}_v}$. Since γ is pricing based, $\Pr_{(v,x)\sim\gamma}[x \notin \text{tieset}_v] = 0$. More formally, when sampling v, x from γ, with probability 1

$$v(x) - \pi(x) = \psi(v) \geq v(x') - \pi(x')$$

for all $x' \in X$. This further implies that $q_{\gamma,X'}(\tilde{x}) = 0$ for all $\tilde{x} \notin X'$ and X' such that $\Pr_{v\sim D}[X' = \text{tieset}_v] > 0$, as otherwise,

$$\begin{aligned}
\Pr_{(v,x)\sim\gamma}[x \notin \text{tieset}_v] &\geq \Pr_{(v,x)\sim\gamma}\left[x \notin \text{tieset}_v, \text{tieset}_v = X'\right] \\
&\geq \Pr_{(v,x)\sim\gamma}\left[x \notin \text{tieset}_v \mid \text{tieset}_v = X'\right] \cdot \Pr_{(v,x)\sim\gamma}\left[\text{tieset}_v = X'\right] \\
&\geq \Pr_{(v,x)\sim\gamma}\left[x = \tilde{x} \mid \text{tieset}_v = X'\right] \cdot \Pr_{(v,x)\sim\gamma}\left[\text{tieset}_v = X'\right] && \textit{(Since } \tilde{x} \notin X'\textit{)} \\
&= q_{\gamma,X'}(\tilde{x}) \cdot \Pr_{(v,x)\sim\gamma}\left[\text{tieset}_v = X'\right] > 0
\end{aligned}$$

Therefore, $\Pr_{v\sim D}[A(v) \notin \text{tieset}_v] = 0$, which in turn implies

$$\begin{aligned}
\mathbb{E}_{(v,x)\sim\gamma}[v(x)] &= \mathbb{E}_{(v,x)\sim\gamma}\left[\max_{x'} (v(x') - \pi(x'))\right] \\
&= \mathbb{E}_{v\sim D}\left[\max_{x'} (v(x') - \pi(x'))\right] \\
&= \mathbb{E}_{v\sim D}[A(v)],
\end{aligned}$$

[12] We do not use the notation $x_1, \ldots, x_n$ here to avoid confusion as the agents were numbered based on their arrival order. The vector $\mathbf{x}$ simply denotes an assignment of all agents.

where the first equality follows from the fact that γ is pricing based, the second equality follows from the fact that the marginal of γ on V is D, and the third equality follows from the fact that $\Pr_{v\sim D}[A(v) \notin \text{tieset}_v] = 0$.

Finally, the marginal of γ on X is $A(D)$ because

$$\begin{aligned}
\Pr_{v\sim D}[A(v) = x'] &= \sum_{X'} \Pr_{v\sim D}[\text{tieset}_v = X'] \Pr_{v\sim D}[A(v) = x' \mid \text{tieset}_v = X'] \\
&= \sum_{X'} \Pr_{v\sim D}[\text{tieset}_v = X'] \, q_{\gamma,X'}(x') && \textit{(Definition of } A(v)\textit{)} \\
&= \sum_{X'} \Pr_{v\sim D}[\text{tieset}_v = X'] \Pr_{(v,x)\sim\gamma}[x = x' \mid \text{tieset}_v = X'] && \textit{(Definition of } q_{\gamma,X'}\textit{)} \\
&= \sum_{X'} \Pr_{(v,x)\sim\gamma}[x = x', \text{tieset}_v = X'] \\
&= \Pr_{(v,x)\sim\gamma}[x = x'] .
\end{aligned}$$

□