

Posted Pricing and Prophet Inequalities with Inaccurate Priors

PAUL DÜTTING, London School of Economics, UK

THOMAS KESSELHEIM, University of Bonn, Germany

In posted pricing, one defines prices for items (or other outcomes), buyers arrive in some order and take their most preferred bundle among the remaining items. Over the last years, our understanding of such mechanisms has improved considerably. The standard assumption is that the mechanism has exact knowledge of probability distribution the buyers' valuations are drawn from. The prices are then set based on this knowledge.

We examine to what extent existing results and techniques are robust to inaccurate prior beliefs. That is, the prices are chosen with respect to similar but different probability distributions. We focus on the question of welfare maximization. We consider all standard distance measures on probability distributions, and derive tight bounds on the welfare guarantees that can be derived for all standard techniques in the various metrics.

CCS Concepts: • **Theory of computation** → **Online algorithms**; **Computational pricing and auctions**; • **Applied computing** → **Economics**.

Additional Key Words and Phrases: posted pricing, prophet inequalities

ACM Reference Format:

Paul Dütting and Thomas Kesselheim. 2019. Posted Pricing and Prophet Inequalities with Inaccurate Priors. In *ACM EC '19: ACM Conference on Economics and Computation (EC '19)*, June 24–28, 2019, Phoenix, AZ, USA. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3328526.3329576>

1 INTRODUCTION

Posted-price mechanisms and prophet inequalities have seen a surge of interest over the past few years. Posted-price mechanisms are probably the most frequently used economic mechanisms for selling goods or services in practice. They are “simple,” truthful, near-optimal mechanisms. Prophet inequalities in turn are the most important mathematical tool for analyzing posted-price mechanisms, but they also provide a less pessimistic framework for analyzing online algorithms than the classic competitive analysis framework by assuming that the input is drawn from known distributions rather than being fully adversarial.

The simplest problem for which posted-price mechanisms and prophet inequalities have been studied is the problem of allocating a single item. Here n agents arrive one-by-one with a value v_i drawn from a known distribution, and upon arrival of an agent it must be decided whether this agent should get the item or not.

A canonical generalization of this problem is to matroid feasibility structures, in which multiple agents can be accepted as long as they form an independent set in a given matroid. Another generalization is to combinatorial auctions, in which multiple items can be allocated and each agent has a valuation function, mapping bundles of items to values.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EC '19, June 24–28, 2019, Phoenix, AZ, USA

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6792-9/19/06...\$15.00

<https://doi.org/10.1145/3328526.3329576>

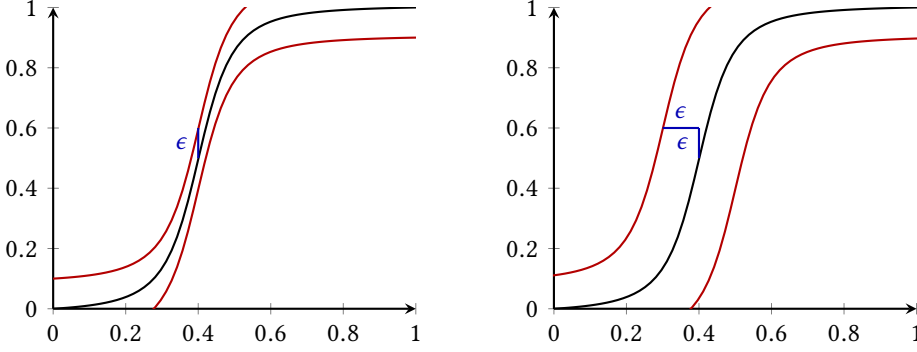


Fig. 1. Kolmogorov metric (on the left) vs. Lévy metric (on the right)

In all of these settings, the optimal online algorithm is a *posted-price mechanism*, which prices all allocations that are still feasible and lets the arriving agent choose a utility maximizing allocation. These optimal prices can be found via *backward induction*.

Computing this optimal strategy, however, may require exponential space or solving a NP-hard optimization problem, it generally requires “dynamic pricing,” and it can be hard to quantify how good this strategy is in comparison to the offline optimum.

Alternative, simpler strategies have been proposed that only approximate the offline optimum. They also serve as constructive ways to establish *prophet inequalities*: The expected value of the chosen allocation provides an α -approximation to the expected optimal value in hindsight. Prophet inequalities can be price-based in which case they show that the corresponding posted-price mechanism obtains an α -approximation of the optimal offline value.

Over the past few years two main techniques for establishing price-based prophet inequalities have been developed, *quantile pricing*, where decisions are made such that probabilities of certain events have certain values, and *balanced prices*, where prices are set high enough so that they partially offset the value lost due to allocation decisions but also low enough so that agents could receive high enough utility from what remains unallocated.

The standard assumption in the literature is that the priors $\mathcal{D}_1, \dots, \mathcal{D}_n$ from which the valuations are drawn are *known precisely*. In almost all practically relevant applications this seems like a rather strong and unrealistic assumption. In this paper, we therefore examine how *robust* these different techniques are when only *inaccurate* priors $\tilde{\mathcal{D}}_1, \dots, \tilde{\mathcal{D}}_n$ are known. Specifically, we will assume that prices are defined based on $\tilde{\mathcal{D}}_1, \dots, \tilde{\mathcal{D}}_n$ but the agents’ values are still drawn from $\mathcal{D}_1, \dots, \mathcal{D}_n$, where each $\tilde{\mathcal{D}}_i$ is ϵ -close to \mathcal{D}_i in a probability metric.

1.1 The Challenges

We illustrate the challenges in developing such results by reviewing the standard techniques for the single-item setting. There are two standard metrics for this setting. Both are defined on the cumulative distribution functions. The *Kolmogorov metric* allows the cumulative distribution functions to be shifted vertically by $\pm\epsilon$. The *Lévy metric* allows the cumulative distribution function to be shifted vertically and horizontally by $\pm\epsilon$. See Figure 1 for a visualization, and Section 3 for formal definitions.

Distances in the Lévy metric are smaller than in the Kolmogorov metric. So assuming that two distributions are ϵ -close in the Kolmogorov metric is a stronger assumption than assuming that they are ϵ -close in the Lévy metric.

Since probability metrics are additive, they naturally lead to additive error terms. We therefore assume that valuations are normalized to $[0, 1]$.

Optimal Static Price. A natural approach is to choose the best static price. This, however, can be arbitrarily bad if distributions are only close in Lévy metric. To see this consider a setting with $n \geq 2$ i.i.d. agents. If the inaccurate priors are that valuations are distributed uniformly on $[1 - \epsilon, 1]$, then the best static price will be some $\tilde{p} \in (1 - \epsilon, 1)$. If the values are actually distributed uniformly on $[1 - 2\epsilon, 1 - \epsilon]$, then this price achieves no welfare at all.

Quantile Pricing. The original prophet inequality proof of Samuel-Cahn [19] sets a single price such that with probability exactly $1/2$ the item is sold. This can be shown to achieve $1/2$ of the maximum expected value on accurate priors.

Like the best optimal static price this approach is not robust if the inaccurate priors are only close in the Lévy metric. A shift of values by ϵ can change the probability of selling drastically; it can even drop to 0. (See Proposition 4.12 in Section 4.3).

Balanced Prices. The canonical balanced prices mechanism for this setting is the mechanism by Kleinberg and Weinberg [15]. It sets a single price, and uses this price for all agents. The price that it sets is $1/2$ the maximum expected value. Setting this price can be shown to recover $1/2$ of the maximum expected value provided that priors are accurate.

If the inaccurate priors are ϵ -close in Kolmogorov or Lévy metric, then the expected maximum value differs by at most $2n\epsilon$ (Lemma 4.10 in Section 4.2). Hence the prices \tilde{p} and p computed on the inaccurate priors and the accurate priors will be within $n\epsilon$ of each other.

Prior work has shown that in this case the expected value that one achieves by setting \hat{p} instead of p is by at most an additive $O(n\epsilon)$ term lower [11]. In case of multiple items, however, it is not difficult to see that prices can be *far apart* when the values can be shifted. (See Example 4.7 in Section 4.2.)

Backward Induction. Defining a sequence of prices that maximize the expected value can be done recursively. For the last agent, the price will be zero; for agent $i < n$ it is exactly the expected value extracted from agents $i + 1, \dots, n$. Here, it is arguably less obvious what to expect. How do errors propagate and accumulate? Are there cascading effects?

1.2 Our Results

We consider (essentially all) standard distance measures on probability distributions (see Section 3), and derive tight bounds on the welfare guarantees that can be derived for all standard techniques in the various metrics.

Lévy, Prokhorov, and Wasserstein Metric. Our first set of results concerns the Lévy metric and similar metrics for multi-parameter settings.

We show that the optimal policy via backward induction for the inaccurate priors $\tilde{\mathcal{D}}$ only loses an additive $O(n\epsilon)$ term compared to the expected welfare achieved by the optimal policy for the actual distributions \mathcal{D} (Theorem 4.1 in Section 4). The crux is to first compare the expected welfare that these two optimal policies achieve on their respective distributions $\tilde{\mathcal{D}}$ and \mathcal{D} . In a second step, we relate the expected welfare that the optimal policy for the inaccurate priors achieves on the actual distributions \mathcal{D} to that on the inaccurate priors $\tilde{\mathcal{D}}$. For both steps we make use of the existence of a coupling which follows from Strassen's Theorem [20].

We then show that the prophet-inequality guarantee of *balanced prices* only loses an additive $O(n\epsilon)$ term (Theorem 4.8 in Section 4). This strengthens the results in prior work by showing that proximity in price space is not necessary. Prices are balanced if they are high enough and low

enough with respect to suitable benchmarks. As we show these benchmarks are close in expectation. For this we again use the existence of the aforementioned couplings.

The additive term of $O(n\epsilon)$ in both these results is optimal, even under the much stronger assumption that the distributions are close in the total variation distance (Proposition 4.5 and Proposition 4.11 in Section 4). In particular, this shows that the error has to increase as the number of agents increases, but only linearly so. In other words, there is no cascading effect.

For *prices defined on quantiles* no such guarantees hold. We show that they can be arbitrarily bad, even for arbitrarily close distributions.

Kolmogorov Metric and Total Variation Distance. We also show results for probability metrics that only allow shifts in probability space but not in value space.

We first consider single-parameter settings and the Kolmogorov metric, and show a result that applies to all approaches reviewed above. Namely, we prove that for *all price-based approaches* a policy's expected welfare differs by at most an additive $O(n\epsilon)$ term, when evaluated on the actual or inaccurate priors (Theorem 5.1 in Section 5). This in particular implies that prophet-inequality type guarantees are preserved. This bound is tight and generally does not hold for approaches that are not based on prices (Proposition 5.2 and Example 5.3 in Section 5).

For general multi-parameter settings and the total variation distance we show that *all policies* only lose an additive $O(n^2\epsilon)$ term (Theorem 6.1 in Section 6), and this bound is tight even for price-based policies (Proposition 6.2 in Section 6).

1.3 Related Work

Quantile Pricing and Balanced Prices. Apart from the first two papers on prophet inequalities by Krengel and Sucheston [16, 17], essentially all subsequent work on prophet inequalities and posted prices either follows the quantile approach or the balanced prices approach. The canonical example for the quantile approach is [19]. Extensions to combinatorial settings that follow this approach include [1, 2, 7, 9, 12]. The prototype for the balanced prices approach is [15]. Other important examples of this approach include [11] and [10].

Sample-Based Approaches. A related but different approach is to assume that one is only given sample access to the priors. This direction is pursued for welfare in [4, 8, 21] and for revenue in [5].

A possible connection is that one approach in this setting would be to use samples to approximately learn the true underlying distributions, and to then apply a known technique to the empirical distributions which will be close to the actual distributions in some metric.

The existing approaches, however, use very different techniques. They usually work with only a single sample from each distribution and crucially rely on the fact that the samples and the valuations come from exactly the same distributions.

Connection to Robust Optimization. An alternative approach to our approach is to follow the perspective of robust optimization. Here, one does not know the instance exactly but one is given an *uncertainty set* of possible instances. The goal is to find a max-min optimal solution.

In the context of monopoly pricing with a single buyer this has been done by Bergemann and Schlag [6], who assume that the designer is given an inaccurate prior $\tilde{\mathcal{D}}$ and an $\epsilon > 0$, and seeks to design a mechanism that maximizes the minimal revenue over all possible actual priors \mathcal{D} that are ϵ -close in the Prokhorov metric to $\tilde{\mathcal{D}}$.

An important difference between this and our approach is that this approach requires ϵ to be known to the mechanism designer.

1.4 Future Work

We believe that this paper could be the starting point for a very fruitful research agenda that examines posted pricing and prophet inequalities with inaccurate priors.

A first direction would be to extend our “unknown uncertainty set” approach to revenue maximization. A second direction would be to design mechanisms that are max-min optimal for both welfare and revenue maximization.

2 PRELIMINARIES

The Problem. We consider general *allocation problems* with n agents. We are looking for an allocation (x_1, \dots, x_n) with the meaning that agent i is assigned $x_i \in X_i$. The space of all feasible such allocations is denoted by $\mathcal{F} \subseteq X_1 \times \dots \times X_n$, $\mathcal{F} \neq \emptyset$. For example, we may consider the problem of allocating bundles of indivisible items to agents, as in a combinatorial auction. In this case, there will be a set of items M . Any agent can be allocated a bundle of items, so $X_i = 2^M$ for all i . The feasible allocations \mathcal{F} is the set of all S_1, \dots, S_n such that S_i and $S_{i'}$ are disjoint for $i \neq i'$. Another example is a matroid structure in which $X_i = \{0, 1\}$ for all i and \mathcal{F} is the set of all characteristic vectors of independent sets in the matroid.

Each agent i has a *valuation* v_i . The valuation is a function $v_i : X_i \rightarrow \mathbb{R}_{\geq 0}$, mapping the respective assignment to a real value. Our goal is to maximize social welfare. That is, we would like to find an allocation $x \in \mathcal{F}$ such that $\sum_{i=1}^n v_i(x_i)$ is maximized. We denote by $\text{OPT}(v) = \max_{x \in \mathcal{F}} \sum_{i=1}^n v_i(x_i)$ the optimal social welfare on valuation profile v .

We assume to be in a setting of *incomplete information*. The valuation v_i of agent i is drawn from a *known* probability distribution \mathcal{D}_i , so $v_i \sim \mathcal{D}_i$. These draws are independent but the distributions are not necessarily identical. We let \mathcal{D} denote the respective induced product distribution over valuation profiles.

We are interested in finding an allocation *online*: The agents arrive one after the other in order $1, \dots, n$. In the i -th step, the assignment for agent i has to be determined only depending on v_1, \dots, v_i . It turns out that, since valuations are independent, nothing can be gained by basing this decision on the actual valuations v_1, \dots, v_{i-1} of previous agents as opposed to only their allocations x_1, \dots, x_{i-1} . This motivates the study of policies in which the allocation decision for agent i only depends on v_i and the allocations x_1, \dots, x_{i-1} made to the previous agents.

Formally, a *policy* is a function π , which maps agent i 's value and the allocation to agents $1, \dots, i-1$ to an allocation for agent i . That is, the allocation is derived by $x_i = \pi(v_i, x_1, \dots, x_{i-1})$. Given any distribution \mathcal{D} , we write $W_{\mathcal{D}}(\pi)$ for the expected social welfare that is achieved by policy π . The easiest way to define $W_{\mathcal{D}}(\pi)$ formally is through the following recursion. Write $W_{\mathcal{D}}(\pi, x_1, \dots, x_{i-1})$ for the expected social welfare of policy π if the first $i-1$ agents are fixed to x_1, \dots, x_{i-1} but not counted in social welfare. This way, $W_{\mathcal{D}}(\pi, x_1, \dots, x_n) = 0$ for all x_1, \dots, x_n and furthermore

$$W_{\mathcal{D}}(\pi, x_1, \dots, x_{i-1}) = \mathbb{E}_{v_i \sim \mathcal{D}_i} [v_i(\pi(v_i, x_1, \dots, x_{i-1})) + W_{\mathcal{D}}(\pi, x_1, \dots, x_{i-1}, \pi(v_i, x_1, \dots, x_{i-1}))] .$$

Backward Induction. An *optimal policy* $\pi_{\mathcal{D}}^*$ for distribution \mathcal{D} is a policy π that maximizes $W_{\mathcal{D}}(\pi)$. The optimal policy can be found via *backward induction* because

$$W_{\mathcal{D}}(\pi_{\mathcal{D}}^*, x_1, \dots, x_{i-1}) = \mathbb{E}_{v_i \sim \mathcal{D}} \left[\max_{x_i \in \mathcal{F}_{x_1, \dots, x_{i-1}}} v_i(x_i) + W_{\mathcal{D}}(\pi_{\mathcal{D}}^*, x_1, \dots, x_{i-1}, x_i) \right] ,$$

where $\mathcal{F}_{x_1, \dots, x_{i-1}} = \{(x'_1, \dots, x'_n) \in \mathcal{F} \mid x'_1 = x_1, \dots, x'_{i-1} = x_{i-1}\}$ is the set of all feasible allocations that complete the partial allocation x_1, \dots, x_{i-1} .

Depending on the problem at hand, backward induction may also require solving a computationally hard problem as well as exponential space.

Posted Pricing. *Posted-price mechanisms* approach each agent i in turn, and offer agent i a menu of prices $p_i(x_i \mid x_1, \dots, x_{i-1})$ for each $x_i \in X_i$, and let agent i choose an allocation x_i that maximizes her utility $v_i(x_i) - p_i(x_i \mid x_1, \dots, x_{i-1})$, with ties broken arbitrarily. We require that $p_i(x_i \mid x_1, \dots, x_{i-1}) = \infty$ for any x_i that is infeasible given x_1, \dots, x_{i-1} , so that the allocations chosen by the agents form a feasible allocation.

Every posted-price mechanism induces a policy, but not every policy can be implemented through a posted-price mechanism. Posted-price mechanisms are interesting in their own right, and they have the advantage that they are *truthful* mechanisms.

It turns out that the optimal policy defined via backward induction always has a correspondence in posted prices (see the full version of this paper for details).

Prophet Inequalities. For a class of distributions over valuation functions, a *prophet inequality* asserts the existence of a family of policies that for each distribution in the class achieves at least an $\alpha \in [0, 1]$ fraction of the expected optimal social welfare. That is, for every distribution \mathcal{D} in the class, there is a policy $\pi_{\mathcal{D}}$ such that

$$W_{\mathcal{D}}(\pi_{\mathcal{D}}) \geq \alpha \cdot \mathbb{E}_{v \sim \mathcal{D}} [\text{OPT}(v)].$$

While for each distribution \mathcal{D} the optimal policy $\pi_{\mathcal{D}}^*$ can be found via backward induction, it is often hard to relate the expected social welfare $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*)$ achieved by this policy to the optimal expected social welfare $\mathbb{E}_{v \sim \mathcal{D}} [\text{OPT}(v)]$. Prophet inequalities are therefore often established via simpler, suboptimal policies.

3 PROBABILITY METRICS

Previous work assumes that decisions can be made based on knowledge of the distribution \mathcal{D} . In an abstract sense, this means that an algorithm knows the distribution \mathcal{D} and computes a policy $\pi_{\mathcal{D}}$ based on knowledge of \mathcal{D} such that the welfare $W_{\mathcal{D}}(\pi_{\mathcal{D}})$ is high. We will consider the case where the algorithm does not actually have access to the distribution \mathcal{D} but only to a similar one $\tilde{\mathcal{D}}$. That is, the algorithm computes a policy $\pi_{\tilde{\mathcal{D}}}$ based on knowledge of $\tilde{\mathcal{D}}$ and we are interested in the welfare $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}})$ that this policy achieves on \mathcal{D} .

In order to make quantitative statements about $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}})$, we first have to formalize what we mean by “ \mathcal{D} and $\tilde{\mathcal{D}}$ are similar”. To this end, we will consider metrics on probability distributions. There are numerous standard ways to define such a metric (see, e.g., [13]). Each of these defines a distance measure on any pair of probability distributions \mathcal{D}_i and $\tilde{\mathcal{D}}_i$. See Figure 2 for an overview and how the various metrics relate to each other.

Recall that we defined probability distribution \mathcal{D}_i to be over functions $v_i: X_i \rightarrow \mathbb{R}_{\geq 0}$. Denoting the set of such functions by V_i , the probability distribution \mathcal{D}_i is defined by the probabilities that it assigns to all Borel sets $S \subseteq V_i$. For a Borel set $S \subseteq V_i$, we write $\mathcal{D}_i(S)$ for the probability of the event that the distribution returns a $v_i \in S$. Some of the distance measures for distributions also require a metric to be defined on V_i . For concreteness, we will consider the metric induced by the ℓ_{∞} -norm. That is, $d(v_i, v'_i) = \|v_i - v'_i\|_{\infty} = \sup_{x_i \in X_i} |v_i(x_i) - v'_i(x_i)|$.

Since all standard distance measures d on probability distributions are additive, requiring that $d(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i naturally leads to additive guarantees of the form $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}) \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}) - f(n, \epsilon)$, where $f: \mathbb{N} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is some function in n and ϵ . We will therefore assume that valuations are normalized such that $v_i: X_i \rightarrow [0, 1]$ for all i .

3.1 Single Parameter

The first two probability metrics are defined for single-parameter settings. In a single-parameter setting $X_i = [0, 1]$ and $v_i(x_i) = v_i^* \cdot x_i$ for some parameter $v_i^* \in [0, 1]$. In such a setting any

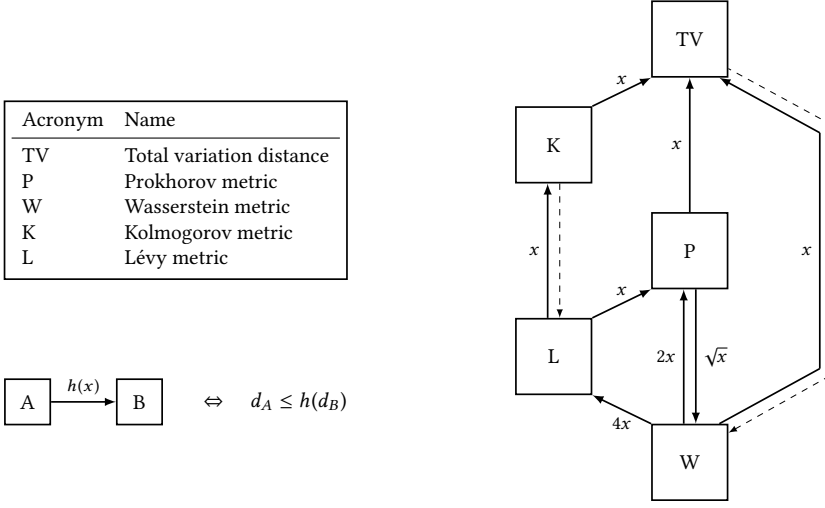


Fig. 2. Relationships among probability metrics (see, e.g., [13]). Dashed lines indicate bounds that rely on further parameters not considered here.

probability distribution \mathcal{D}_i on V_i can be identified with the corresponding cumulative distribution function F_i on $[0, 1]$ via $F_i(z) = \Pr_{v_i \sim \mathcal{D}_i} [v_i(1) \leq z]$.

Similarity of \mathcal{D}_i and $\tilde{\mathcal{D}}_i$ can now be defined by the similarity of the respective cumulative distribution functions F_i and \tilde{F}_i .

Kolmogorov Metric. The *Kolmogorov metric* bounds the maximum absolute difference between $F_i(z)$ and $\tilde{F}_i(z)$ over $z \in [0, 1]$. Formally,

$$d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i) = \sup_{z \in [0, 1]} |F_i(z) - \tilde{F}_i(z)| .$$

So requiring that $d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ corresponds to taking F_i , moving it up and down by ϵ , and requiring that \tilde{F}_i is fully contained in the resulting band.

Lévy Metric. The *Lévy metric* also compares F_i to \tilde{F}_i , but allows vertical and horizontal shifts. Namely,

$$d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i) = \inf\{\epsilon : \tilde{F}_i(z - \epsilon) - \epsilon \leq F_i(z) \leq \tilde{F}_i(z + \epsilon) + \epsilon \text{ for all } z \in [0, 1]\} .$$

Relationships. The Kolmogorov metric is a stronger requirement than the Lévy Metric. In particular, whenever $d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for some ϵ then also $d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ and, thus, $d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$ [14, p. 34]. The reverse direction is generally unbounded. If, for example, $F_i(z) = 1$ for $z \geq 0$ and 0 otherwise but $\tilde{F}_i(z) = 1$ for $z \geq \epsilon$ and 0 otherwise, then $d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i) = \epsilon$ but $d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i) = 1$. However, there are still bounds relying on further parameters (see, e.g., [18, p. 43]).

3.2 Beyond Single Parameter

The remaining distance measures apply to the general case, and, e.g., capture the problem of allocating a set of indivisible items to agents, as in a combinatorial auction. Recall that in this particular case we would have $X_i = 2^M$, where M is the set of indivisible items, and \mathcal{D}_i would be a distribution over set functions $v_i : 2^M \rightarrow [0, 1]$.

Total Variation Distance. The total variation distance of \mathcal{D}_i and $\tilde{\mathcal{D}}_i$, is defined by

$$d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) = \sup_{S \subseteq V_i, S \text{ is a Borel set}} |\mathcal{D}_i(S) - \tilde{\mathcal{D}}_i(S)| . \quad (1)$$

The total variation distance has a coupling characterization, which is an implication of Strassen's Theorem.

LEMMA 3.1 ([20]). *For any two distributions $\mathcal{D}_i, \tilde{\mathcal{D}}_i$, we have $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ if and only if there is a joint distribution \mathcal{D}_i^* over pairs (v_i, \tilde{v}_i) such that the marginal distributions of v_i and \tilde{v}_i are \mathcal{D}_i and $\tilde{\mathcal{D}}_i$ respectively and it holds that $\Pr[v_i \neq \tilde{v}_i] < \epsilon$.*

Prokhorov Metric. To define the Prokhorov metric, we first need to define for a set $S \subseteq V_i$ and $\epsilon > 0$ its ϵ -neighborhood $S^\epsilon \subseteq V_i$ by

$$S^\epsilon = \{v_i \in V_i \mid \exists v'_i \in S, d(v_i, v'_i) < \epsilon\} .$$

The Prokhorov metric of \mathcal{D}_i and $\tilde{\mathcal{D}}_i$, denoted by $d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$, is the infimum over all ϵ such that for all Borel sets $S \subseteq V_i$ we have

$$\mathcal{D}_i(S) \leq \tilde{\mathcal{D}}_i(S^\epsilon) + \epsilon . \quad (2)$$

The Prokhorov metric is particularly relevant in robust statistics because a sequence of probability measures weakly converges if and only if the Prokhorov distance to the limit vanishes. Due to Strassen's Theorem, it also has a coupling characterization.

LEMMA 3.2 ([20]). *For any two distributions $\mathcal{D}_i, \tilde{\mathcal{D}}_i$, we have $d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ if and only if there is a joint distribution \mathcal{D}_i^* over pairs (v_i, \tilde{v}_i) such that the marginal distributions of v_i and \tilde{v}_i are \mathcal{D}_i and $\tilde{\mathcal{D}}_i$ respectively and it holds that $\Pr[\|v_i - \tilde{v}_i\|_\infty > \epsilon] < \epsilon$.*

Wasserstein Metric. The Wasserstein metric has many equivalent formulations. For example, for discrete distributions it has the interpretation of an *earth mover's distance*. The question is to transform one distribution into the other by shifting masses in a way that minimizes the sum of weighted distances.

For our purposes, however, it is most convenient to define it by the existence of a coupling. We say that the Wasserstein metric between \mathcal{D}_i and $\tilde{\mathcal{D}}_i$, denoted by $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$ is the infimum over all ϵ for which there is a joint distribution \mathcal{D}_i^* over pairs (v_i, \tilde{v}_i) such that the marginal distributions of v_i and \tilde{v}_i are \mathcal{D}_i and $\tilde{\mathcal{D}}_i$ respectively and it holds that $\mathbb{E}[\|v_i - \tilde{v}_i\|_\infty] < \epsilon$.

Relationships. The total variation distance when specialized to distributions on $[0, 1]$ requires that condition (1) is satisfied for all Borel sets, while the Kolmogorov metric only requires this condition to be satisfied for intervals that start at zero. Therefore, $d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$. Similarly, the Prokhorov metric when specialized to $[0, 1]$ would require condition (2) to be satisfied for all Borel sets, while the Lévy metric only requires this condition to be satisfied for intervals that start at zero. So, $d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$. Finally, it is possible to relate the Wasserstein metric when specialized to distributions on $[0, 1]$ to the Lévy metric, and show that $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq 4d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$ (see the full version of this paper for details).

From the coupling characterizations, it is clear that the total variation distance is always lower-bounded by the Prokhorov distance, i.e., $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \geq d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$, because the requirement on the coupling is stronger. For this reason, assuming that \mathcal{D}_i is close to $\tilde{\mathcal{D}}_i$ in the Prokhorov distance is a weaker assumption than in the total variation distance.

Requiring that two distributions are ϵ -close in the Wasserstein metric is a weaker requirement than requiring that they are ϵ -close in the total variation distance, i.e., $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$,

because for the coupling for which $\Pr[v_i \neq \tilde{v}_i] < \epsilon$ we have $\mathbf{E}[\|v_i - \tilde{v}_i\|_\infty] \leq \Pr[v_i = \tilde{v}_i] \cdot 0 + \Pr[v_i \neq \tilde{v}_i] \cdot 1 = \Pr[v_i \neq \tilde{v}_i] < \epsilon$.

Also note that if two distributions are ϵ -close in the Prokhorov metric, then the respective coupling fulfills $\mathbf{E}[\|v_i - \tilde{v}_i\|_\infty] \leq \Pr[\|v_i - \tilde{v}_i\|_\infty < \epsilon] + \epsilon < 2\epsilon$. Therefore, $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq 2d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$. Finally, by Markov's inequality, $\mathbf{E}[\|v_i - \tilde{v}_i\|_\infty] \leq \epsilon$ implies $\Pr[\|v_i - \tilde{v}_i\|_\infty > \sqrt{\epsilon}] < \sqrt{\epsilon}$, and therefore, $d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \sqrt{d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i)}$.

Further Metrics. There are a few other standard distance measures studied in the literature (cf. [13]), such as Hellinger distance or Kullback-Leibler divergence, which are all polynomially lower-bounded in the total variation distance. For instance, the Hellinger distance is lower bounded by the total variation distance while the Kullback-Leibler distance is lower bounded by two times the total variation distance squared.

We will see that for the strongest possible results, it suffices to require that the distributions are close in total variation distance (see Section 6).

4 LÉVY, PROKHOROV, AND WASSERSTEIN METRIC

We begin with the weakest set of assumptions—that the inaccurate priors are close in the Lévy, Prokhorov, or Wasserstein metric—and show that the optimal policy obtained via backward induction and balanced prices are robust, while quantile pricing is not.

More specifically, we assume that each of the inaccurate priors $\tilde{\mathcal{D}}_i$ is ϵ -close to the accurate prior \mathcal{D}_i , and show the following:

(i) For the optimal policy obtained via backward induction we show that the expected social welfare that it obtains is at least $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) - O(n\epsilon)$, and that the additive error term of $O(n\epsilon)$ is asymptotically sharp even under the most stringent assumption that each of the distributions is ϵ -close in total variation distance.

(ii) For balanced prices we show that, if the approach guarantee an expected social welfare of at least $\gamma \cdot \mathbf{E}[\text{OPT}(v)]$, $\gamma \in [0, 1]$ on the accurate priors, then on the inaccurate priors it will obtain an expected social welfare of at least $\gamma \cdot \mathbf{E}[\text{OPT}(v)] - O(n\epsilon)$. Again, this additive error term of $O(n\epsilon)$ is asymptotically tight even if we assume that the inaccurate priors are ϵ -close in total variation distance.

(iii) We show that the expected social welfare obtained by quantile pricing, when tuned to inaccurate priors, can be arbitrarily worse than what the respective approach would yield with accurate priors.

We think that the linear error term in (i) is somewhat surprising as it shows that the errors in each distribution do *not* propagate (and lead to a super-linear error term) as one may have suspected. Our results in (ii) apply even if the balanced prices obtained on the inaccurate priors and those obtained on the accurate priors are *far apart*. They thus strengthen previous robustness results in [11] and [10], which showed an additive error term of $O(n\epsilon)$ when $\|p - \tilde{p}\| \leq \epsilon$.

4.1 Backward Induction

We state and prove our main result in this section. It concerns the backward-induction policy $\pi_{\tilde{\mathcal{D}}}^*$ for inaccurate distributions $\tilde{\mathcal{D}}$, where each $\tilde{\mathcal{D}}_i$ is ϵ -close to the accurate distribution \mathcal{D}_i according to the Wasserstein metric. It shows that the welfare $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*)$ obtained by this policy on the true underlying distributions \mathcal{D} is within $2n\epsilon$ of the (optimal) welfare $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*)$ obtained by the backward-induction policy $\pi_{\mathcal{D}}^*$ for these distributions.

THEOREM 4.1. *If $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ for all i , then $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*) \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) - 2n\epsilon$.*

An immediate implication of this theorem and the fact that $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq 2d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$ and $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq 4d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$ is that the same (up to constant factors) additive error term holds for the Prokhorov and Lévy metric.

COROLLARY 4.2. *If $d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ for all i or $d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ for all i , then $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) \geq W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*) - O(n\epsilon)$.*

Our proof of Theorem 4.1 will make use of the existence of couplings \mathcal{D}_i^* between each pair of distributions \mathcal{D}_i and $\tilde{\mathcal{D}}_i$ such that $\mathbf{E}[\|v_i - \tilde{v}_i\|_\infty] < \epsilon$. Specifically, for each i we will assume that the pair of valuation functions (v_i, \tilde{v}_i) is drawn from distribution \mathcal{D}_i^* . All probabilities and expectations are over these joint distributions.

To prove the theorem we proceed in two steps. First, in Lemma 4.3, we show that $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) - n\epsilon$. That is, we show that the values that the optimal policies for $\tilde{\mathcal{D}}$ and \mathcal{D} achieve on the respective distributions is close. Afterwards, in Lemma 4.4, we show that $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) \geq W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) - n\epsilon$. This means that we evaluate the policy that is optimal for $\tilde{\mathcal{D}}$ on \mathcal{D} but compare it to the value that it achieves on $\tilde{\mathcal{D}}$.

LEMMA 4.3. *If $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ for all i , then $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) - n\epsilon$.*

PROOF. We show that for all $i \in [n+1]$ and all x'_1, \dots, x'_{i-1} ,

$$W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*, x'_1, \dots, x'_{i-1}) - (n-i+1)\epsilon$$

by downward induction on i .

The base case for our induction is $i = n+1$. In this case, we have $W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}) = 0$ and $W_{\tilde{\mathcal{D}}}(\pi, x'_1, \dots, x'_{i-1}) = 0$ for any policy π , so the claim holds.

In the induction step, we consider a fixed $i \in [n+1]$ and x_1, \dots, x_{i-1} and we assume that the statement holds for $i+1$.

Let

$$x_i = \pi_{\mathcal{D}}^*(v_i, x'_1, \dots, x'_{i-1}) \quad \text{and} \quad \tilde{x}_i = \pi_{\tilde{\mathcal{D}}}^*(\tilde{v}_i, x'_1, \dots, x'_{i-1})$$

be the choices of the respective optimal policies given that the allocations for agents $1, \dots, i-1$ are x'_1, \dots, x'_{i-1} . Note that these are defined such that \tilde{x}_i maximizes

$$\tilde{v}_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) .$$

So, therefore

$$\tilde{v}_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \geq \tilde{v}_i(x_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i) .$$

On $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i)$, we can apply the induction hypothesis. Furthermore, $\tilde{v}_i(x_i) \geq v_i(x_i) - \|\tilde{v}_i - v_i\|_\infty$. This combines to

$$\tilde{v}_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \geq v_i(x_i) + W_{\mathcal{D}}(\pi_{\mathcal{D}}^*, x'_1, \dots, x'_{i-1}, x_i) - (n-i)\epsilon - \|\tilde{v}_i - v_i\|_\infty .$$

Taking the expectation of both sides, we get

$$\begin{aligned} W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) &= \mathbf{E} \left[\tilde{v}_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \right] \\ &\geq \mathbf{E} \left[v_i(x_i) + W_{\mathcal{D}}(\pi_{\mathcal{D}}^*, x'_1, \dots, x'_{i-1}, x_i) \right] - (n-i)\epsilon - \mathbf{E} [\|\tilde{v}_i - v_i\|_\infty] \\ &= W_{\mathcal{D}}(\pi_{\mathcal{D}}^*, x'_1, \dots, x'_{i-1}) - (n-i)\epsilon - \mathbf{E} [\|\tilde{v}_i - v_i\|_\infty] . \end{aligned}$$

Hence, by assumption, $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*, x'_1, \dots, x'_{i-1}) - (n-i+1)\epsilon$, as required. \square

LEMMA 4.4. *If $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ for all i , then $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) \geq W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) - n\epsilon$.*

PROOF. As in the proof of the previous lemma, we proceed by induction. This time we show that for all $i \in [n + 1]$ and all x'_1, \dots, x'_{i-1} ,

$$W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) \geq W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) - (n - i + 1)\epsilon$$

by downward induction on i .

Again, the base case is $i = n + 1$. In this case, we have that $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) = 0$ and $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) = 0$, so the claim holds.

For the induction step, consider any fixed i and x'_1, \dots, x'_{i-1} . This time, we define

$$x_i = \pi_{\tilde{\mathcal{D}}}^*(v_i, x'_1, \dots, x'_{i-1}) \quad \text{and} \quad \tilde{x}_i = \pi_{\tilde{\mathcal{D}}}^*(\tilde{v}_i, x'_1, \dots, x'_{i-1})$$

to be the choices of $\pi_{\tilde{\mathcal{D}}}^*$ if the valuation of agent i is v_i or \tilde{v}_i , respectively.

By definition,

$$\begin{aligned} W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) &= \mathbf{E} \left[v_i(x_i) + W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i) \right] \quad \text{and} \\ W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) &= \mathbf{E} \left[\tilde{v}_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \right]. \end{aligned}$$

Furthermore,

$$\tilde{v}_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \leq v_i(\tilde{x}_i) + \|\tilde{v}_i - v_i\|_{\infty} + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i)$$

Recall that x_i maximizes $v_i(x_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i)$ because $\pi_{\tilde{\mathcal{D}}}^*$ is an optimal policy for distribution $\tilde{\mathcal{D}}$, so in particular,

$$v_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \leq v_i(x_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i).$$

Finally, by the induction hypothesis,

$$W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i) \leq W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i) + (n - i)\epsilon.$$

In combination

$$\begin{aligned} W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) &= \mathbf{E} \left[\tilde{v}_i(\tilde{x}_i) + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \right] \\ &\leq \mathbf{E} \left[v_i(\tilde{x}_i) + \|\tilde{v}_i - v_i\|_{\infty} + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, \tilde{x}_i) \right] \\ &\leq \mathbf{E} \left[v_i(x_i) + \|\tilde{v}_i - v_i\|_{\infty} + W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i) \right] \\ &\leq \mathbf{E} \left[v_i(x_i) + \|\tilde{v}_i - v_i\|_{\infty} + W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}, x_i) + (n - i)\epsilon \right] \\ &= W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) + (n - i)\epsilon + \mathbf{E} [\|\tilde{v}_i - v_i\|_{\infty}]. \end{aligned}$$

By assumption, $\mathbf{E} [\|\tilde{v}_i - v_i\|_{\infty}] < \epsilon$. We obtain $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) \leq W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*, x'_1, \dots, x'_{i-1}) + (n - i + 1)\epsilon$ as claimed. \square

PROOF OF THEOREM 4.1. By Lemma 4.4 and Lemma 4.3,

$$W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*) \geq W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) - n\epsilon \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) - 2n\epsilon. \quad \square$$

We conclude by showing that even under the more stringent assumption of closeness under total variation distance, the additive error term of $O(n\epsilon)$ in Theorem 4.1 and Corollary 4.2 is asymptotically sharp.

PROPOSITION 4.5. *Even for a single item for all $n \geq 2$ and $0 < \epsilon \leq \frac{1}{n}$, there are distributions $\mathcal{D}, \tilde{\mathcal{D}}$ with $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i such that $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}^*) \leq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) - \frac{1}{3}n\epsilon$.*

PROOF. Define distributions \mathcal{D} and $\tilde{\mathcal{D}}$ with $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ as follows. In both distributions the first value is deterministic, and equal to $n\epsilon/3$. In \mathcal{D} the remaining $n-1$ values are equal to 0. In $\tilde{\mathcal{D}}$ they are 1 with probability ϵ and 0 otherwise.

The backward induction policy $\pi_{\tilde{\mathcal{D}}}^*$ for $\tilde{\mathcal{D}}$ accepts the first value, and has $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) = n\epsilon/3$. The backward induction policy $\pi_{\mathcal{D}}^*$ for \mathcal{D} skips the first value and then accepts the first value that is 1, achieving $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) = 1 - (1 - \epsilon)^{n-1} > n\epsilon/3$ but $W_{\tilde{\mathcal{D}}}(\pi_{\mathcal{D}}^*) = 0$. \square

4.2 Balanced Prices

We next consider the technique of setting balanced prices, which leads to posted-price mechanisms that are approximately optimal. We show that social welfare guarantees obtained through this technique for accurate priors $\mathcal{D}_1, \dots, \mathcal{D}_n$ continue to hold up to an additive error of $O(n\epsilon)$ when prices are set based on inaccurate priors $\tilde{\mathcal{D}}_1, \dots, \tilde{\mathcal{D}}_n$ that are ϵ -close to the accurate priors in the Lévy, Prokhorov, or Wasserstein metric. We complement this result with a construction that shows that this error term is asymptotically tight.

While our results hold in the most general version of the balanced prices framework developed in [10], we will present them for problems that can be cast as constrained item allocation problems for ease of exposition.

In such a problem we are given a set of items M , agents can receive subsets of items so that $X_i = 2^M$ for all i , and feasible allocations $\mathcal{F} \subseteq X_1 \times \dots \times X_n$ consist of disjoint allocations of items that may be required to satisfy additional constraints. Note that both combinatorial auctions and matroids can be captured this way.

To define balanced prices we need the following definitions. For $x \in \mathcal{F}$ let \mathcal{F}_x be the set of all $x' \in \mathcal{F}$ such that x' can still be carried out *after* x . Formally, this means that the two allocations x, x' are disjoint, i.e., $x_i \cap x'_i = \emptyset$ for all i , and their component-wise union is feasible, i.e., $x \cup x' := (x_1 \cup x'_1, \dots, x_n \cup x'_n) \in \mathcal{F}$. For a given valuation profile v , define $\text{OPT}(v) = \arg \max_{x \in \mathcal{F}} \sum_i v_i(x_i)$ and $v(\text{OPT}(v)) = \max_{x \in \mathcal{F}} \sum_i v_i(x_i)$. Similarly, for a given valuation profile v and feasible allocation $x \in \mathcal{F}$, define $\text{OPT}(v \mid x) = \arg \max_{x' \in \mathcal{F}_x} \sum_i v_i(x'_i)$ and $v(\text{OPT}(v \mid x)) = \max_{x' \in \mathcal{F}_x} \sum_i v_i(x'_i)$.

Definition 4.6 ([10]). Let $\alpha > 0, \beta > 0$. Given valuation profile v , a pricing rule p^v is (α, β) -balanced if for all $x \in \mathcal{F}$ and all $x' \in \mathcal{F}_x$,

- (a) $\sum_i p_i^v(x_i \mid x_1, \dots, x_{i-1}) \geq \frac{1}{\alpha} \cdot (v(\text{OPT}(v)) - v(\text{OPT}(v \mid x)))$, and
- (b) $\sum_i p_i^v(x'_i \mid x_1, \dots, x_{i-1}) \leq \beta \cdot v(\text{OPT}(v \mid x))$.

A collection of pricing rules $(p^v)_{v \in V}$ is (α, β) -balanced if for each $v \in V = V_1 \times \dots \times V_n$ the pricing rule p^v is (α, β) -balanced.

Dütting et al. [10] show that if $(p^v)_{v \in V}$ is (α, β) -balanced then the posted-prices mechanism $\pi_{\mathcal{D}}$ based on $\delta \cdot p^{\mathcal{D}}$ defined by $p_i^{\mathcal{D}}(x_i \mid y) = \mathbf{E}_{v \sim \mathcal{D}} [p_i^{v'}(x_i \mid y)]$ and $\delta = \frac{\alpha}{1+\alpha\beta}$ fulfills,

$$W_{\mathcal{D}}(\pi_{\mathcal{D}}) \geq \frac{1}{1 + \alpha\beta} \cdot \mathbf{E}_{v \sim \mathcal{D}} \left[\max_{x \in \mathcal{F}} \sum_{i=1}^n v_i(x_i) \right].$$

Dütting et al. [10] also observe that balanced prices are robust in that if the pricing rule p is perturbed to some \hat{p} with $\|p - \hat{p}\|_{\infty} \leq \epsilon$, then the welfare guarantee degrades by at most an additive $O(n\epsilon)$ term.

We start with the observation that, even if $d_p(\tilde{\mathcal{D}}_i, \mathcal{D}_i) \leq \epsilon$ for all i , balanced prices \tilde{p} that are computed based on $\tilde{\mathcal{D}}_1, \dots, \tilde{\mathcal{D}}_n$ can be far apart from the respective prices p that are computed based on $\mathcal{D}_1, \dots, \mathcal{D}_n$.

Example 4.7. Consider the posted-price mechanism of [11] for two agents, two items setting in which the agents $i \in \{1, 2\}$ have unit demand valuations, i.e., $v_i(\emptyset) = 0$, $v_i(\{j\}) = v_{i,j}$ for $j \in \{1, 2\}$, and $v_i(\{1, 2\}) = \max\{v_{i,1}, v_{i,2}\}$.

Specialized to this scenario the mechanism of Feldman et al. computes (1,1)-balanced prices by computing a social welfare maximizing allocation, and setting the (unscaled) price of item j to the valuation $v_{i,j}$ of the agent i that is assigned this item.

Consider the following full information settings \mathcal{D} and $\tilde{\mathcal{D}}$ with $d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i . In \mathcal{D} agent 1 has a value of $v_1(\{1\}) = 1$ for item 1 and value $v_1(\{2\}) = 0$ for item 2 and agent 2 has a value of $v_2(\{1\}) = 2 - \epsilon/2$ for item 1 and a value of $v_2(\{2\}) = 1$ for item 2. In $\tilde{\mathcal{D}}$ the values of agent 1 are unchanged so that $\tilde{v}_1(\{j\}) = v_1(\{j\})$ for $j \in \{1, 2\}$ and agent 2 has a value of $\tilde{v}_2(\{1\}) = 2 + \epsilon/2$ for item 1 and a value of $\tilde{v}_2(\{2\}) = 1$ for item 2.

For \mathcal{D} the welfare maximizing allocation gives item 1 to agent 1 and item 2 to agent 2, resulting in (unscaled) prices of $p(\{1\}) = 1$ and $p(\{2\}) = 1$. In $\tilde{\mathcal{D}}$, in contrast, the welfare maximizing allocation gives item 1 to agent 2 and item 2 to agent 1, resulting in (unscaled) prices of $\tilde{p}(\{1\}) = 2 + \epsilon/2$ and $\tilde{p}(\{2\}) = 0$. Hence $\|p - \tilde{p}\|_\infty = 1 + \epsilon/2$.

Note that this example uses the fact that prices are discontinuous in the valuations. Indeed, as we demonstrate in the full version of this paper, under a continuity assumption prices are guaranteed to be close. Nonetheless, despite this potential non-proximity in price space we have:

THEOREM 4.8. *Let $\pi_{\tilde{\mathcal{D}}}$ be a posted-price mechanism based on an (α, β) -balanced pricing rule using priors $\tilde{\mathcal{D}}$. If $d_W(\tilde{\mathcal{D}}_i, \mathcal{D}_i) < \epsilon$ for all i , then*

$$W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}) \geq \frac{1}{1 + \alpha\beta} \cdot \mathbb{E}_{v \sim \mathcal{D}} \left[\max_{x \in \mathcal{F}} \sum_{i=1}^n v_i(x_i) \right] - 2n\epsilon .$$

Since $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq 2d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$ and $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq 4d_L(\mathcal{D}_i, \tilde{\mathcal{D}}_i)$ we obtain the following corollary for the Prokhorov and Lévy metric.

COROLLARY 4.9. *If $d_P(\tilde{\mathcal{D}}_i, \mathcal{D}_i) < \epsilon$ for all i or $d_L(\tilde{\mathcal{D}}_i, \mathcal{D}_i) < \epsilon$ for all i , then the posted-price mechanism $\pi_{\tilde{\mathcal{D}}}$ defined based on $\tilde{\mathcal{D}}$ fulfills*

$$W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}) \geq \frac{1}{1 + \alpha\beta} \cdot \mathbb{E}_{v \sim \mathcal{D}} \left[\max_{x \in \mathcal{F}} \sum_{i=1}^n v_i(x_i) \right] - O(n\epsilon) .$$

To prove Theorem 4.8, we follow generally the same argument as in [10], which uses Property (b) to show a lower bound on the utility and Property (a) to show a lower bound on the revenue and then combines these into a welfare guarantee.

The difference is that the distribution for valuations \mathcal{D} differs from the distribution used to define the prices $\tilde{\mathcal{D}}$. As we have already seen this can make prices differ significantly. However, this is not true for the right-hand sides of Property (a) and Property (b) in expectation. Indeed, we can show the following lemma.

LEMMA 4.10. *Consider any $x \in \mathcal{F}$. If $d_W(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ for all i , then for $v \sim \mathcal{D}$, $\tilde{v} \sim \tilde{\mathcal{D}}$ we have*

$$\mathbb{E} [\tilde{v}(\text{OPT}(\tilde{v} \mid x))] \geq \mathbb{E} [v(\text{OPT}(v \mid x))] - n\epsilon .$$

PROOF. Note that $\tilde{v}_i(x'_i) \geq v_i(x'_i) - \|v_i - \tilde{v}_i\|_\infty$ for all x'_i . So also

$$\tilde{v}(\text{OPT}(\tilde{v} \mid x)) = \max_{x' \in \mathcal{F}_x} \sum_{i=1}^n \tilde{v}_i(x'_i) \geq \max_{x' \in \mathcal{F}_x} \sum_{i=1}^n (v_i(x'_i) - \|v_i - \tilde{v}_i\|_\infty) = v(\text{OPT}(v \mid x)) - \sum_{i=1}^n \|v_i - \tilde{v}_i\|_\infty .$$

As we have $\mathbf{E} [\|v_i - \tilde{v}_i\|_\infty] < \epsilon$ by assumption, linearity of expectation gives us

$$\mathbf{E} [\tilde{v}(\text{OPT}(\tilde{v} \mid x))] \geq \mathbf{E} [v(\text{OPT}(v \mid x))] - \sum_{i=1}^n \mathbf{E} [\|v_i - \tilde{v}_i\|_\infty] \geq \mathbf{E} [v(\text{OPT}(v \mid x))] - n\epsilon \quad \square$$

Based on this lemma, we can now follow the same steps as in [10] with only minor modifications. We present the full argument here for completeness.

PROOF OF THEOREM 4.8. The argument in [10] considers three valuation profiles v, v' , and \tilde{v} . The first profile is the one the mechanism is run on, the second one is used to construct potential deviations, and the final one defines the prices. In the original proof these are identically distributed, we instead consider $v, v' \sim \mathcal{D}$ and $\tilde{v} \sim \tilde{\mathcal{D}}$.

We will write $x(v)$ for the allocation returned by the posted-price mechanism on input valuation profile v , and $x_{[i-1]}(v)$ for $x_1(v), \dots, x_{i-1}(v)$. We use \tilde{p} as a shorthand for $p^{\tilde{\mathcal{D}}}$.

Utility bound: We obtain a lower bound on the expected utility of an agent as follows. We sample valuations $v' \sim \mathcal{D}$. Agent i now considers buying $\text{OPT}_i((v_i, v'_{-i}) \mid x(v'_i, v'_{-i}))$ at price $\delta \cdot \tilde{p}_i(\text{OPT}_i((v_i, v'_{-i}) \mid x(v'_i, v'_{-i})) \mid x_{[i-1]}(v))$. Taking expectations and exploiting that $x_{[i-1]}(v)$ does not depend on v_i we obtain

$$\begin{aligned} \mathbf{E}_v [u_i(v)] &\geq \mathbf{E}_{v, v'} \left[v_i \left(\text{OPT}_i((v_i, v'_{-i}) \mid x(v'_i, v'_{-i})) \right) - \delta \cdot \tilde{p}_i \left(\text{OPT}_i((v_i, v'_{-i}) \mid x(v'_i, v'_{-i})) \mid x_{[i-1]}(v) \right) \right] \\ &= \mathbf{E}_{v, v'} \left[v'_i \left(\text{OPT}_i(v' \mid x(v)) \right) - \delta \cdot \tilde{p}_i \left(\text{OPT}_i(v' \mid x(v)) \mid x_{[i-1]}(v) \right) \right]. \end{aligned}$$

Summing the previous inequality over all agents we get

$$\begin{aligned} \mathbf{E}_v \left[\sum_{i=1}^n u_i(v) \right] &\geq \mathbf{E}_{v, v'} \left[\sum_{i=1}^n v'_i \left(\text{OPT}_i(v' \mid x(v)) \right) \right] - \mathbf{E}_{v, v'} \left[\sum_{i=1}^n \delta \cdot \tilde{p}_i \left(\text{OPT}_i(v' \mid x(v)) \mid x_{[i-1]}(v) \right) \right] \\ &= \mathbf{E}_{v, v'} \left[v' \left(\text{OPT}(v' \mid x(v)) \right) \right] - \mathbf{E}_{v, v'} \left[\sum_{i=1}^n \delta \cdot \tilde{p}_i \left(\text{OPT}_i(v' \mid x(v)) \mid x_{[i-1]}(v) \right) \right]. \end{aligned} \quad (3)$$

We can upper bound the last term in the previous inequality by using Property (b). This gives

$$\sum_{i=1}^n \delta \cdot \tilde{p}_i \left(\text{OPT}_i(v' \mid x(v)) \mid x_{[i-1]}(v) \right) \leq \delta \beta \cdot \mathbf{E}_{\tilde{v}} \left[\tilde{v} \left(\text{OPT}(\tilde{v} \mid x(v)) \right) \right]$$

pointwise for any v and v' , and therefore also

$$\mathbf{E}_{v, v'} \left[\sum_{i=1}^n \delta \cdot \tilde{p}_i \left(\text{OPT}_i(v' \mid x(v)) \mid x_{[i-1]}(v) \right) \right] \leq \delta \beta \cdot \mathbf{E}_{v, \tilde{v}} \left[\tilde{v} \left(\text{OPT}(\tilde{v} \mid x(v)) \right) \right]. \quad (4)$$

As v' and \tilde{v} are not identically distributed, the last step in the original argument cannot be applied. Instead, combining (3) with (4), we get

$$\mathbf{E}_v \left[\sum_{i=1}^n u_i(v) \right] \geq \mathbf{E}_{v, v'} [v'(\text{OPT}(v' \mid x(v)))] - \delta \beta \cdot \mathbf{E}_{v, \tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v} \mid x(v)))] \quad (5)$$

Revenue bound: For the revenue bound we can proceed in the original proof. In particular, by applying Property (a),

$$\sum_{i=1}^n \delta \cdot \tilde{p}_i(x_i(v) \mid x_{[i-1]}(v)) = \delta \cdot \sum_{i=1}^n \mathbf{E}_{\tilde{v}} [p_i^{\tilde{v}}(x_i(v) \mid x_{[i-1]}(v))]$$

$$\geq \frac{\delta}{\alpha} \cdot \mathbf{E}_{\tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v})) - \tilde{v}(\text{OPT}(\tilde{v} \mid x(v)))].$$

Taking expectation over v this shows

$$\mathbf{E}_v \left[\sum_{i=1}^n \delta \cdot \tilde{p}_i(x_i(v) \mid x_{[i-1]}(v)) \right] \geq \frac{\delta}{\alpha} \cdot \mathbf{E}_{\tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v}))] - \frac{\delta}{\alpha} \cdot \mathbf{E}_{v, \tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v} \mid x(v)))]. \quad (6)$$

Combination: We conclude by combining the utility and revenue bounds (5) and (6). Using $\delta = \frac{\alpha}{1+\alpha\beta}$ we get

$$\begin{aligned} W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}) &= \mathbf{E}_v \left[\sum_{i=1}^n v_i(x_i(v)) \right] = \mathbf{E}_v \left[\sum_{i=1}^n u_i(v) \right] + \mathbf{E}_v \left[\sum_{i=1}^n \delta \cdot \tilde{p}_i(x_i(v) \mid x_{[i-1]}(v)) \right] \\ &\geq \mathbf{E}_{v, v'} [v'(\text{OPT}(v' \mid x(v)))] - \delta\beta \cdot \mathbf{E}_{v, \tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v} \mid x(v)))] \\ &\quad + \frac{\delta}{\alpha} \cdot \mathbf{E}_{\tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v}))] - \frac{\delta}{\alpha} \cdot \mathbf{E}_{v, \tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v} \mid x(v)))] \\ &\geq \frac{1}{1+\alpha\beta} \mathbf{E}_{\tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v}))] + \mathbf{E}_{v, v'} [v'(\text{OPT}(v' \mid x(v)))] - \mathbf{E}_{v, \tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v} \mid x(v)))] . \end{aligned}$$

By Lemma 4.10, we have that

$$\mathbf{E}_{\tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v}))] \geq \mathbf{E}_v [v(\text{OPT}(v))] - n\epsilon .$$

As the Wasserstein distance is symmetric, we can also swap the roles of $\tilde{\mathcal{D}}$ and \mathcal{D} in Lemma 4.10. We therefore also have for any fixed v

$$\mathbf{E}_{\tilde{v}} [\tilde{v}(\text{OPT}(\tilde{v} \mid x(v)))] \leq \mathbf{E}_{v'} [v'(\text{OPT}(v' \mid x(v)))] + n\epsilon .$$

In combination, this gives us

$$W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}) \geq \frac{1}{1+\alpha\beta} \cdot \mathbf{E}_v [\text{OPT}(v)] - 2n\epsilon . \quad \square$$

The error bound of $\Theta(n\epsilon)$ is again tight, even under the strongest assumption of ϵ -closeness under total variation distance. This is true even for the mechanism of Kleinberg and Weinberg [15] for a single item, the prototypical example of balanced prices, which sets the price to 1/2 of the expected optimal social welfare.

PROPOSITION 4.11. *Even for a single item and the policy π of Kleinberg and Weinberg [15], for all $n \geq 2$, $0 < \epsilon \leq \frac{1}{n}$, and $\delta > 0$ there are distributions $\mathcal{D}, \tilde{\mathcal{D}}$ with $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i such that $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}) \leq (1/(1+\alpha\beta)) \cdot \mathbf{E}_{v \sim \mathcal{D}} [\max_{x \in \mathcal{F}} \sum_i v_i(x_i)] - \frac{1}{6}n\epsilon + \delta$.*

PROOF. Consider the single-item setting with $n \geq 2$ agents from Proposition 4.5 with the roles of \mathcal{D} and $\tilde{\mathcal{D}}$ interchanged, and with the value of the first agent set to δ . That is, in both \mathcal{D} and $\tilde{\mathcal{D}}$ the value of the first agent is deterministic, and equal to δ . In \mathcal{D} the value of each of the remaining $n-1$ agents is 1 with probability ϵ , and 0 otherwise. In $\tilde{\mathcal{D}}$ the value of the remaining $n-1$ agents is deterministic and equal to 0.

Now consider the mechanism of Kleinberg and Weinberg [15] for this setting, which computes $(1, 1)$ -balanced prices. This mechanism sets the (scaled) price of the item to 1/2 of the expected optimal social welfare. So the price computed on $\tilde{\mathcal{D}}$ is $\delta/2$, which achieves the same expected social welfare $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}})$ and $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}})$ on $\tilde{\mathcal{D}}$ and \mathcal{D} , namely δ . The expected optimal social welfare on \mathcal{D} , in contrast, is $\mathbf{E}_{v \sim \mathcal{D}} [\max_i v_i] \geq 1 - (1-\epsilon)^{n-1} > n\epsilon/3$ and so $1/2 \cdot \mathbf{E}_{v \sim \mathcal{D}} [\max_i v_i] > n\epsilon/6$. \square

4.3 Quantile Pricing

Besides approaches defining balanced prices, the other standard approach to define approximate policies is to make decisions such that probabilities of certain events have certain values. We subsume this class of approaches under the term *quantile pricing*. Examples are the single-item prophet inequality proof by Samuel-Cahn [19], which sets a price such that the sequence is stopped with probability exactly $\frac{1}{2}$. Other examples include the recent results on single-item i.i.d. prophet inequalities [1, 9], the prophet inequalities of [2] and [3] for unit-demand agents, and more generally approaches based on online contention resolution schemes [12].

Unfortunately, as it turns out, quantile pricing is usually not robust to deviations to neighboring distributions with respect to the Prokhorov metric.

PROPOSITION 4.12. *Even for the single-item setting and the policy π of Samuel-Cahn [19], for every $n \geq 2$ and every $\epsilon > 0$ there exists distributions \mathcal{D} and $\tilde{\mathcal{D}}$ with $d_P(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i such that $W_{\mathcal{D}}(\pi_{\tilde{\mathcal{D}}}) = 0$ but $\mathbb{E}_{v \sim \mathcal{D}} [\max_i v_i] \geq W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) \geq 1 - 2\epsilon$.*

PROOF. We consider only a single item and only a single buyer. Given any $\epsilon > 0$, let $\mathcal{D} = \text{Uniform}[1 - 2\epsilon, 1 - \epsilon]$ and $\tilde{\mathcal{D}} = \text{Uniform}[1 - \epsilon, 1]$. Observe that $d_P(\mathcal{D}, \tilde{\mathcal{D}}) \leq \epsilon$. The policy π that ensures that the sequence is stopped with probability $1/2$ on distribution $\tilde{\mathcal{D}}$ uses threshold $\tau = 1 - \frac{\epsilon}{2}$. So, $W_{\mathcal{D}}(\pi) = 0$ because the sequence never gets stopped. This is in contrast to $\mathbb{E}_{v \sim \mathcal{D}} [\max_i v_i] \geq 1 - 2\epsilon$ and $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) \geq 1 - 2\epsilon$. \square

5 KOLMOGOROV METRIC

In this section, we strengthen our assumptions and derive a much stronger guarantee. We consider single-parameter settings. That is, $X_i = [0, 1]$ and $v_i(x_i) = v_i^* \cdot x_i$ for all i . That is, each agent's valuation is completely described by one real number v_i^* . Note that the single-item setting but also matroid constraints can be captured this way.

We consider policies induced by a posted-price mechanism. We show that any such policy is robust with respect to deviations in the Kolmogorov metric by showing the following theorem.

THEOREM 5.1. *Consider a single-parameter setting and a policy π induced by a posted-price mechanism. If $d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i , then $W_{\mathcal{D}}(\pi) \geq W_{\tilde{\mathcal{D}}}(\pi) - n\epsilon$.*

An implication of the previous theorem is that guarantees with respect to the expected optimal value and the optimal policy via backward induction are preserved.

Indeed, if π guarantees that $W_{\tilde{\mathcal{D}}}(\pi) \geq \alpha \mathbb{E}_{\tilde{v} \sim \tilde{\mathcal{D}}} [\text{OPT}(\tilde{v})]$ or $W_{\tilde{\mathcal{D}}}(\pi) \geq \alpha W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*)$ for some $0 < \alpha \leq 1$, then also $W_{\mathcal{D}}(\pi) \geq \alpha \mathbb{E}_{v \sim \mathcal{D}} [\text{OPT}(v)] - 5n\epsilon$ by Lemma 4.10 or $W_{\mathcal{D}}(\pi) \geq \alpha W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) - 5n\epsilon$ by Lemma 4.3 respectively.

PROOF OF THEOREM 5.1. We prove this theorem by downward induction and show that for all $i \in [n + 1]$ and all x'_1, \dots, x'_{i-1} we have

$$W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}) \geq W_{\tilde{\mathcal{D}}}(\pi, x'_1, \dots, x'_{i-1}) - (n - i + 1)\epsilon.$$

For $i = n + 1$, we have

$$W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}) = W_{\tilde{\mathcal{D}}}(\pi, x'_1, \dots, x'_{i-1}) = 0$$

for all x'_1, \dots, x'_{i-1} , so the claim holds.

Let us now consider an $i < n + 1$. For a fixed choice of x'_1, \dots, x'_{i-1} , define $x_i(v_i) = p_i(1 \mid x'_1, \dots, x'_{i-1})$. By this definition

$$\begin{aligned} & v_i(\pi(v_i, x'_1, \dots, x'_{i-1})) + W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}, \pi(v_i, x'_1, \dots, x'_{i-1})) \\ &= v_i(x_i(v_i)) + W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}, x_i(v_i)) \end{aligned}$$

and therefore also

$$W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}) = \mathbf{E}_{v_i \sim \mathcal{D}} [v_i(x_i(v_i)) + W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}, x_i(v_i))] .$$

By induction hypothesis, we have $W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}, x_i(v_i)) \geq W_{\tilde{\mathcal{D}}}(\pi, x'_1, \dots, x'_{i-1}, x_i(v_i)) - (n-i)\epsilon$ because x'_1, \dots, x'_{i-1} and v_i are fixed.

Due to the fact that π is defined by posted prices, we have that $x_i(v_i) \geq x''_i$ if and only if

$$v_i(x_i(v_i)) - p_i(x_i(v_i) \mid x'_1, \dots, x'_{i-1}) \geq v_i(x''_i) - p_i(x''_i \mid x'_1, \dots, x'_{i-1}) .$$

Because v_i is non-decreasing, $x_i(\cdot)$ has to be a non-decreasing function. Therefore, we have $v_i(x_i(v_i)) > t$ if $v_i(1) \in I$, where I is an interval of the form $[t', \infty)$ or (t', ∞) . We can use $d_K(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ to get

$$\begin{aligned} \mathbf{Pr}_{v_i \sim \mathcal{D}_i} [v_i(x_i(v_i)) \geq t] &= \mathbf{Pr}_{v_i \sim \mathcal{D}_i} [v_i(1) \in I] \\ &\geq \mathbf{Pr}_{\tilde{v}_i \sim \tilde{\mathcal{D}}_i} [v_i(1) \in I] - \epsilon = \mathbf{Pr}_{\tilde{v}_i \sim \tilde{\mathcal{D}}_i} [v_i(x_i(v_i)) \geq t] - \epsilon . \end{aligned}$$

This implies

$$\begin{aligned} \mathbf{E}_{v_i \sim \mathcal{D}_i} [v_i(x_i(v_i))] &= \int_0^1 \mathbf{Pr}_{v_i \sim \mathcal{D}_i} [v_i(x_i(v_i)) \geq t] dt \\ &\geq \int_0^1 (\mathbf{Pr}_{\tilde{v}_i \sim \tilde{\mathcal{D}}_i} [v_i(x_i(v_i)) \geq t] - \epsilon) dt = \mathbf{E}_{\tilde{v}_i \sim \tilde{\mathcal{D}}_i} [v_i(x_i(v_i))] - \epsilon . \end{aligned}$$

In combination

$$W_{\mathcal{D}}(\pi, x'_1, \dots, x'_{i-1}) \geq W_{\tilde{\mathcal{D}}}(\pi, x'_1, \dots, x'_{i-1}) - (n-i+1)\epsilon . \quad \square$$

The error term in Theorem 5.1 is tight:

PROPOSITION 5.2. *Even for a single item for all $n \geq 2$ and $0 < \epsilon \leq \frac{1}{n}$, there are distributions $\mathcal{D}, \tilde{\mathcal{D}}$ with $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i such that $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) \leq W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) - \frac{1}{3}n\epsilon$.*

PROOF. We define the distributions \mathcal{D} and $\tilde{\mathcal{D}}$ with $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ as follows. In \mathcal{D} all n values are equal to 0. In $\tilde{\mathcal{D}}$ they are 1 with probability ϵ and 0 otherwise.

The policy $\pi_{\tilde{\mathcal{D}}}^*$ accepts the first agent with value 1. Therefore, $W_{\tilde{\mathcal{D}}}(\pi_{\tilde{\mathcal{D}}}^*) = 1 - (1 - \epsilon)^n > n\epsilon/3$. However, clearly $W_{\mathcal{D}}(\pi_{\mathcal{D}}^*) = 0$ regardless of the policy. \square

We conclude this section by observing that Theorem 5.1 does not hold for arbitrary policies π , which are not defined by thresholds.

Example 5.3. Consider the setting of only a single agent and a single item. For any $k \in \mathbb{N}$, define $\mathcal{D}_1 = \text{Uniform}[\frac{1}{k}, \frac{2}{k}, \dots, 1]$ and $\tilde{\mathcal{D}}_1 = \text{Uniform}[0, 1]$. Note that for all $t \in \mathbb{R}$ we have $\mathbf{Pr}_{v_1 \sim \mathcal{D}_1} [v_1 \leq t] \leq \mathbf{Pr}_{v_1 \sim \tilde{\mathcal{D}}_1} [v_1 \leq t] \leq \mathbf{Pr}_{v_1 \sim \mathcal{D}_1} [v_1 \leq t] + \frac{1}{k}$, so $d_K(\mathcal{D}_1, \tilde{\mathcal{D}}_1) \leq \frac{1}{k}$. Now consider the policy π , which accepts any agent unless their value is a multiple of $\frac{1}{k}$. We observe that $W_{\mathcal{D}}(\pi) = 0$ but $W_{\tilde{\mathcal{D}}}(\pi) = \frac{1}{2}$, regardless of the value of k .

6 TOTAL VARIATION DISTANCE

In this section, we discuss the consequences of the distributions being close with respect to the total variation distance. As we will show, every policy is robust against such deviations.

THEOREM 6.1. *Consider any policy π . If $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ for all i , then $W_{\mathcal{D}}(\pi) \geq W_{\tilde{\mathcal{D}}}(\pi) - n^2\epsilon$.*

As in the case of Theorem 5.1, Theorem 6.1 implies that guarantees with respect to the expected optimal value or the optimal policy via backward induction are preserved up to an additive $O(n^2\epsilon)$ term.

PROOF OF THEOREM 6.1. We use the fact that the total variation distance has a coupling characterization (Lemma 3.1). We have $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) < \epsilon$ if and only if there is a joint distribution over pairs (v_i, \tilde{v}_i) such that v_i is distributed according to \mathcal{D}_i and \tilde{v}_i is distributed according to $\tilde{\mathcal{D}}_i$ and $\Pr[v_i \neq \tilde{v}_i] < \epsilon$.

By drawing from each of these joint distributions once, we get a probability distribution over valuation profiles such that, by union bound, $\Pr[\exists i : v_i \neq \tilde{v}_i] < n\epsilon$. That is, running π in parallel on v_1, \dots, v_n and $\tilde{v}_1, \dots, \tilde{v}_n$, with probability at least $1 - n\epsilon$ these executions are the same. Otherwise, we use the trivial upper bound of n on the difference. \square

The bound in Theorem 6.1 is again tight:

PROPOSITION 6.2. *There exists a policy π and a single-parameter setting such that for every $n \geq 2$, $0 \leq \epsilon < 1/n$, and $\delta > 0$ there are distributions \mathcal{D} and $\tilde{\mathcal{D}}$ with $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) \leq \epsilon$ for all i for which $W_{\mathcal{D}}(\pi) \leq W_{\tilde{\mathcal{D}}}(\pi) - \Omega(\epsilon n^2) + n\epsilon\delta$.*

PROOF. Consider the following binary single-parameter setting: Either $\{1, \dots, \frac{n}{2}\}$ or $\{\frac{n}{2}, \dots, n\}$ can be accepted simultaneously as well as any subset. Let π be the greedy policy that accepts the first agent who has a strictly positive value and following agents if their value is positive and it would be feasible to add the agent.

We now define the distributions \mathcal{D} and $\tilde{\mathcal{D}}$. Let $\delta > 0$. For $i \leq \frac{n}{2}$, we choose $\tilde{v}_i(1)$ to be 0 with probability 1. Define $\tilde{v}_i(1) = \delta$ with probability ϵ and 0 otherwise. By this choice $d_{TV}(\mathcal{D}_i, \tilde{\mathcal{D}}_i) = \epsilon$ regardless of δ . Furthermore, for $i > \frac{n}{2}$, we define $v_i(1)$ and $\tilde{v}_i(1)$ to be 1 with probability 1.

Observe that $W_{\mathcal{D}}(\pi) \leq \frac{n}{2} - \Omega(\epsilon n^2) + n\epsilon\delta$ but $W_{\tilde{\mathcal{D}}}(\pi) = \frac{n}{2}$ for any $\delta > 0$ if $\epsilon \leq \frac{1}{n}$. \square

REFERENCES

- [1] Melika Abolhassani, Soheil Ehsani, Hossein Esfandiari, MohammadTaghi Hajiaghayi, Robert D. Kleinberg, and Brendan Lucier. 2017. Beating $1 - 1/e$ for Ordered Prophets. In *Proceedings of the 49th ACM Symposium on Theory of Computing (STOC'17)*. 61–71.
- [2] Saeed Alaei. 2014. Bayesian Combinatorial Auctions: Expanding Single Buyer Mechanisms to Many Buyers. *SIAM J. Comput.* 43, 2 (2014), 930–972.
- [3] Saeed Alaei, MohammadTaghi Hajiaghayi, and Vahid Liaghat. 2012. Online Prophet-Inequality Matching with Applications to Ad Allocation. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC'12)*. 18–35.
- [4] Pablo Daniel Azar, Robert Kleinberg, and S. Matthew Weinberg. 2014. Prophet Inequalities with Limited Information. In *Proceedings of the 25th ACM-SIAM Symposium on Discrete Algorithms (SODA'14)*. 1358–1377.
- [5] Moshe Babaioff, Liad Blumrosen, Shaddin Dughmi, and Yaron Singer. 2017. Posting Prices with Unknown Distributions. *ACM Transactions on Economics and Computation* 5, 2 (2017), 13:1–13:20.
- [6] Dirk Bergemann and Karl H. Schlag. 2011. Robust Monopoly Pricing. *Journal of Economic Theory* 146, 6 (2011), 2527–2543.
- [7] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. 2010. Multi-Parameter Mechanism Design and Sequential Posted Pricing. In *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC'10)*. 311–320.
- [8] José Correa, Paul Dütting, Felix A. Fischer, and Kevin Schewior. 2019. Prophet Inequalities for I.I.D. Random Variables from an Unknown Distribution. In *Proceedings of the 20th ACM Conference on Economics and Computation (EC'19)*. Forthcoming.
- [9] José Correa, Patricio Foncea, Ruben Hoeksma, Tim Oosterwijk, and Tjark Vredeveld. 2017. Posted Price Mechanisms for a Random Stream of Customers. In *Proceedings of the 18th ACM Conference on Economics and Computation (EC'17)*. 169–186.
- [10] Paul Dütting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. 2017. Prophet Inequalities Made Easy: Stochastic Optimization by Pricing Non-Stochastic Inputs. In *Proceedings of the 58th IEEE Symposium on Foundations of Computer Science (FOCS'17)*. 540–551.

- [11] Michal Feldman, Nick Gravin, and Brendan Lucier. 2015. Combinatorial Auctions via Posted Prices. In *Proceedings of the 26th ACM-SIAM Symposium on Discrete Algorithms (SODA'15)*. 123–135.
- [12] Moran Feldman, Ola Svensson, and Rico Zenklusen. 2016. Online Contention Resolution Schemes. In *Proceedings of the 27th ACM-SIAM Symposium on Discrete Algorithms (SODA'16)*. 1014–1033.
- [13] A. L. Gibbs and F. E. Su. 2002. On Choosing and Bounding Probability Metrics. *International Statistical Review* 70, 3 (2002), 419–435.
- [14] P. J. Huber. 1981. *Robust Statistics*. John Wiley & Sons, New York City, NY, USA.
- [15] Robert Kleinberg and S. Matthew Weinberg. 2012. Matroid Prophet Inequalities. In *Proceedings of the 44th ACM Symposium on Theory of Computing Conference (STOC'12)*. 123–136.
- [16] U. Krengel and L. Sucheston. 1977. Semiamarts and Finite Values. *Bull. Amer. Math. Soc.* 83 (1977), 745–747.
- [17] U. Krengel and L. Sucheston. 1978. On Semiamarts, Amarts, and Processes with Finite Value. *Advances in Probability and Related Topics* 4 (1978), 197–266.
- [18] V. V. Petrov. 1995. *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*. Oxford University Press, New York City, NY, USA.
- [19] E. Samuel-Cahn. 1984. Comparison of Threshold Stop Rules and Maximum for Independent Nonnegative Random Variables. *Annals of Probability* 12 (1984), 1213–1216. Issue 4.
- [20] V Strassen. 1965. The Existence of Probability Measures with Given Marginals. *Annals of Mathematical Statistics* 36 (1965), 423–439.
- [21] Jack Wang. 2018. The Prophet Inequality Can Be Solved Optimally with a Single Set of Samples. *arXiv: 1812.10563* (2018). Retrieved from: <http://arxiv.org/abs/1812.10563>.