Computing Simple Mechanisms: Lift-and-Round over Marginal Reduced Forms

Yang Cai*
Computer Science Department
Yale University
yang.cai@yale.edu

Argyris Oikonomou†
Computer Science Department
Yale University
argyris.oikonomou@yale.edu

Mingfei Zhao‡
Google Research
mingfei@google.com

April 12, 2022

Abstract

We study revenue maximization in multi-item multi-bidder auctions under the natural *item-independence* assumption – a classical problem in Multi-Dimensional Bayesian Mechanism Design. One of the biggest challenges in this area is developing algorithms to compute (approximately) optimal mechanisms that are not brute-force in the size of the bidder type space, which is usually exponential in the number of items in multi-item auctions. Unfortunately, such algorithms were only known for basic settings of our problem when bidders have unit-demand [CHMS10b, CMS15] or additive valuations [Yao15].

In this paper, we significantly improve the previous results and design the first algorithm that runs in time *polynomial in the number of items and the number of bidders* to compute mechanisms that are $O(1)$-approximations to the optimal revenue when bidders have XOS valuations, resolving the open problem raised in [CM16, CZ17]. Moreover, the computed mechanism has a simple structure: It is either a posted price mechanism or a two-part tariff mechanism. As a corollary of our result, we show how to compute an approximately optimal and simple mechanism efficiently using *only sample access* to the bidders' value distributions. Our algorithm builds on two innovations that allow us to search over the space of mechanisms efficiently: (i) a new type of succinct representation of mechanisms – the *marginal reduced forms*, and (ii) a novel *Lift-and-Round procedure* that concavifies the problem.

*Supported by a Sloan Foundation Research Fellowship and the NSF Award CCF-1942583 (CAREER).
†Supported by a Sloan Foundation Research Fellowship and the NSF Award CCF-1942583 (CAREER).
‡Work done in part while the author was studying at Yale University, supported by the NSF Award CCF-1942583 (CAREER).

1 Introduction

Revenue-maximization in multi-item auctions has been recognized as a central problem in Economics and more recently in Computer Science. While Myerson's celebrated work showed that a simple mechanism is optimal in single-item settings [Mye81], the optimal multi-item mechanism is known to be prohibitively complex and notoriously difficult to characterize even in basic settings. Facing the challenge, a major research effort has been dedicated to understanding the computational complexity for finding an approximately revenue-optimal mechanism in multi-item settings. Despite significant progress, there is still a substantial gap in our understanding of the problem, for example, in the natural and extensively studied *item-independent* setting, first introduced in the influential paper by Chawla, Hartline, and Kleinberg [CHK07].

Formally, the item-independent setting is defined as follows: A seller is selling m heterogeneous items to n bidders, where the i-th bidder's type is drawn independently from an m-dimensional product distribution $D_i = \bigtimes_{j\in[m]} D_{ij}$.[1] We only understand the computational complexity of finding the revenue-optimal mechanism in the item-independent setting for the two most basic valuations: unit-demand and additive valuations. First, we know that finding an exactly optimal mechanism is computationally intractable even for a single bidder with either unit-demand [CDO+15] or additive valuation [DDT14]. Second, there exists a polynomial time algorithm that computes a mechanism whose revenue is at least a constant fraction of the optimal revenue when bidders have unit-demand [CHMS10b, CMS15] or additive valuations [Yao15]. However, unit-demand and additive valuations are only two extremes within a broader class of value functions known as the constrained additive valuations, where the bidder's value is additive subject to a downward-closed feasibility constraint.[2] Furthermore, all constrained additive valuations are contained in an even more general class known as the XOS valuations. Beyond unit-demand and additive valuations, our understanding was limited, and we only knew how to compute an approximately optimal mechanism when bidders are symmetric, i.e., all D_i's are identical [CM16, CZ17]. Finding a polynomial time algorithm for asymmetric bidders was thus raised as a major open problem in both [CM16, CZ17]. In this paper, we resolve this open problem.

Result I: For the item-independent setting with (asymmetric) XOS bidders, there exists an algorithm that computes a Dominant Strategy Incentive Compatible (DSIC) and Individually Rational (IR) mechanism that achieves at least $c \cdot \text{OPT}$ for some absolute constant $c > 0$, where OPT is the optimal revenue achievable by any Bayesian Incentive Compatible (BIC) and IR mechanism. Our algorithm has running time polynomial in $\sum_{i\in[n],j\in[m]} |\mathcal{T}_{ij}|$, where $\mathcal{T}_{ij}$ is the support of D_{ij}. See Theorem 1 for the formal statement.

Computing Approximately Optimal Mechanisms under Structured Distributions. When the bidders' types are drawn from arbitrary distributions, a line of works provide algorithms for finding almost revenue-optimal mechanisms in multi-item settings in time polynomial in the total number of types, i.e., $\sum_{i\in[n]} |\textsc{Supp}(D_i)|$ ($\textsc{Supp}(D_i)$ denotes the support of D_i) [AFH+12, CDW12a, CDW12b, CDW13a, CDW13b, COVZ21]. However, the total number of types could be exponential in the number of items, e.g., there are $\sum_{i\in[n]} \left(\prod_{j\in[m]} |\mathcal{T}_{ij}|\right)$ types in the item-independent case, making these algorithms unsuitable. For unstructured type distributions, such dependence is unavoidable as even describing the distributions requires time $\Omega\left(\sum_{i\in[n]} |\textsc{Supp}(D_i)|\right)$. What if the type distributions are *structured and permit a more succinct description*, e.g., product measures? Arguably, high-dimensional distributions that arise in practice (such as bidders' type distributions in multi-item auctions) are rarely arbitrary, as arbitrary high-dimensional distributions cannot be represented or learned efficiently; see e.g. [DDK18] for a discussion. Indeed, one of the biggest challenges in Bayesian Algorithmic Mechanism Design is designing algorithms to compute (approximately) optimal mechanisms that are not brute-force in the size of the bidder type space when the type distributions are structured. In this paper, we

[1] $[m]$ denotes $\{1, 2, ..., m\}$. D_{ij} is the distribution of bidder i's value for item j. The definition is extended to XOS in Section 2.

[2] A bidder has constrained-additive valuation if the bidder's value for a bundle S is defined as $\max_{V\in 2^S\cap\mathcal{I}} \sum_{j\in V} t_j$, where t_j is the bidder's value for item j, and $\mathcal{I}$ is a downward-closed set system over the items specifying the feasible bundles. Note that constrained-additive valuations contain familiar valuations such as additive, unit-demand, or matroid-rank valuations.

develop computational tools to exploit the item-independence to obtain an exponential speed-up in running time.

Simple vs. Optimal. An additional feature of our algorithm is that the mechanisms computed have a simple structure. It is either a *posted price mechanism* or a *two-part tariff mechanism*. Given the description of the two mechanisms, it is clear that both of them are DSIC and IR.

Rationed Posted Price Mechanism (RPP). There is a price p_{ij} for bidder i to purchase item j. The bidders arrive in some arbitrary order, and each bidder can purchase *at most one* item among the available ones at the given price.[3]

Two-part Tariff Mechanism (TPT). All bidders face the same set of prices $\{p_j\}_{j\in[m]}$. Bidders arrive in some arbitrary order. For each bidder, we show her the available items and the associated price for each item, then ask her to pay an entry fee depending on the bidder's identity and the available items. If the bidder accepts the entry fee, she proceeds to purchase any of the available items at the given prices; if she rejects the entry fee, then she pays nothing and receives nothing.

A recent line of works focus on designing simple and approximately optimal mechanisms [CHK07, CHMS10b, Ala11, HN12, KW12, CH13, BILW14, Yao15, RW15, CDW16, CM16, CZ17]. The main take-away of these results is that in the item-independent setting, there exists a simple mechanism that achieves a constant fraction of the optimal revenue. The most general setting where such a simple $O(1)$-approximation is known is exactly the setting in **Result I**, where bidders have XOS valuations [CZ17]. More specifically, [CZ17] show that there is a RPP or TPT that achieves a constant fraction of the optimal revenue, however their result is purely existential and does not suggest how to compute these simple mechanisms. Our result makes their existential result constructive.

Finally, combining our result with the learnability result for multi-item auctions in [CD17], we can extend our algorithm to the case when we only have sample access to the distributions.

Result II: For constrained-additive bidders, there exists an algorithm that computes a simple, DSIC, and IR mechanism whose revenue is at least $c\cdot \text{OPT} - O(\varepsilon\cdot \text{poly}(n,m))$ for some absolute constant $c>0$ in time polynomial in n, m, and $1/\varepsilon$, given sample access to bidders' type distributions, and assuming each bidder's value for each item lies in $[0,1]$. See Theorem 4 for the formal statement.

1.1 Our Approach and Techniques

Our main technical contribution is *a novel relaxation of the revenue optimization problem that can be solved approximately in polynomial time* and *an accompanying rounding scheme that converts the solution to a simple and approximately optimal mechanism*.[4] Our first step is to replace the objective of revenue with a duality-based benchmark of the revenue proposed in [CZ17]. One can view the new objective as maximizing the virtual welfare, similar to Myerson's elegant solution for the single-item case. The main difference is that, while one can use *a fixed set of virtual valuations for any allocation* in the single-item case, due to the multi-dimensionality of our problem, *the virtual valuations must depend on the allocation, causing the virtual welfare to be a non-concave function in the allocation*. In this paper, we develop algorithmic tools to concavify and approximately optimize the virtual welfare maximization problem. We believe our techniques will be useful to address other similar challenges in Multi-Dimensional Mechanism Design.

More specifically, for every BIC and IR mechanism $\mathcal{M}$ with allocation rule σ and payment rule p, one can choose a set of *dual parameters* $\theta(\sigma)$ based on σ to construct an upper bound $U(\sigma,\theta(\sigma))$ for the revenue of $\mathcal{M}$. We refer to θ as the dual parameters because θ corresponds to a set of "canonical" dual variables, which can be used to derive the virtual valuations via the Cai-Devanur-Weinberg duality framework [CDW16]. The upper

[3] Usually, posted price mechanisms do not restrict the maximum number of items a bidder can buy. We consider a rationed version of posted price mechanism to make the computational task easy.

[4] An influential framework known as the ex-ante relaxation has been widely used in Mechanism Design, but is insufficient for our problem. See Appendix B.2 for a detailed discussion.

bound $U(\sigma, \theta(\sigma))$ is then simply the corresponding virtual welfare. The computational problem is to find an allocation σ that (approximately) maximizes $U(\sigma, \theta(\sigma))$. With such a σ, we could use the result in [CZ17] to convert it to a simple and approximately optimal mechanism. Unfortunately, the function $U(\sigma, \theta(\sigma))$ is highly non-concave in σ,[5] and thus hard to maximize efficiently. See Section 3.1 for a detailed discussion.

LP Relaxation via Lifting. We further relax our objective, i.e., $U(\sigma, \theta(\sigma))$, to obtain a computationally tractable problem. One specific difficulty in optimizing $U(\sigma, \theta(\sigma))$ comes from the fact that $\theta(\sigma)$ is highly non-linear in σ. We address this difficulty in two steps. In the first step of our relaxation, we flip the dependence of σ and θ by relaxing the problem to the following two-stage optimization problem (Figure 1):

- **Stage I:** Maximize $H(\theta)$ subject to some constraints. $H(\theta)$ is the optimal value of the Stage II problem.
- **Stage II:** Maximize an LP over σ with θ-dependent constraints.

This makes the problem much more structured and significantly disentangles the complex dependence between σ and θ. Yet we still do not know how to solve it efficiently. In the second step of our relaxation, we merge the two-stage optimization into a single LP. In particular, we *lift the problem to a higher dimensional space and optimize over joint distributions of the allocation σ and the dual parameters θ via an LP (Figure 3).* Since the number of dual parameters is already exponential in the number of bidders and the number of items, it is too expensive to represent such a joint distribution explicitly. We show it is unnecessary to search over all joint distributions. By leveraging the independence across bidders and items, it suffices for us to consider a set of succinctly representable distributions – the ones whose marginals over the dual parameters are product measures. See Section 3.1 for a more detailed discussion on the development of our relaxation.

"Rounding" any Feasible Solution to a Simple Mechanism. Can we still approximate the optimal solution of the LP relaxation using a simple mechanism? Unfortunately, the result from [CZ17] no longer applies. We provide a *generalization of [CZ17]*, that is, given any feasible solution of our LP relaxation, we can construct in polynomial time a simple mechanism whose revenue is at least a constant fraction of the objective value of the feasible solution (Theorem 3). Our proof (in Appendix C.6) provides several novel ideas to handle the new challenges due to the relaxation, which may be of independent interest.

Marginal Reduced Forms. We deliberately postpone the discussion on how we represent the allocation of a mechanism until now. A widely used succinct representation a mechanism $\mathcal{M}$ is known as the reduced form or the interim allocation rule: $\{r_{ij}(t_i)\}_{i\in[n],j\in[m],t_i\in\times_{j\in[m]}\mathcal{T}_{ij}}$ where $r_{ij}(t_i)$ is the probability for bidder i to receive item j when *her type is* $t_i = (t_{i1}, \dots, t_{im})$ [CDW12a]. Despite being more succinct than the ex-post allocation rule, the reduced form is still too expensive to store in our setting, as its size is exponential in m. A key innovation in our relaxation is the introduction of an even more succinct representation – the *marginal reduced forms* and a *multiplicative approximation* to the polytope of all feasible marginal reduced forms. Although this is a natural concept, to the best of our knowledge, we are the first to introduce and make use of it. The marginal reduced form is represented as : $\left\{w_{ij}(t_{ij})\right\}_{i\in[n],j\in[m],t_{ij}\in\mathcal{T}_{ij}}$, where $w_{ij}(t_{ij})$ is the probability for bidder i to receive item j in $\mathcal{M}$ and *her value for item j is t_{ij}*.[6] Importantly, the size of a marginal reduced form is polynomial in the input size of our problem. As our LP relaxation uses marginal reduced forms as decision variables, it is crucial for us to be able to optimize over the polytope P that contains all *feasible marginal reduced forms*. To the best of our knowledge, P does not have a succinct explicit description or an efficient separation oracle. To overcome the obstacle, we provide *an efficient separation oracle for a different polytope Q that is a multiplicative approximation to P, i.e., $c \cdot P \subseteq Q \subseteq P$ for some absolute constant* $c \in (0, 1)$ (Theorem 2). Using the separation oracle for Q, we can find a c-approximation to the optimum of the LP relaxation efficiently. Note that a sampling technique was developed in [CDW12b] to approximate the polytope of feasible reduced forms. However, their technique only provides an "additive approximation

[5] See Appendix B.1 for an example of the non-concavity of the function.

[6] We refer to $\{w_{ij}(t_{ij})\}_{i\in[n],j\in[m],t_{ij}\in\mathcal{T}_{ij}}$ as the marginal reduced forms as they are the marginals of the reduced forms multiplied by the probability that t_{ij} is bidder i's value for item j, i.e., $\frac{w_{ij}(t_{ij})}{\Pr_{D_{ij}}[t_{ij}]} = \mathbb{E}_{t_{i,-j}\sim\times_{\ell\neq j}D_{i\ell}}\left[r_{ij}\left(t_{ij}, t_{i,-j}\right)\right]$.

to the polytope", which is insufficient for our purpose. Indeed, our multiplicative approximation holds for a wide class of polytopes that frequently appear in Mechanism Design (Theorem 8). We believe our technique has further applications, for example, to convert the additive FPRAS of Cai-Daskalakis-Weinberg [CDW12a, CDW12b, CDW13a, CDW13b] to a multiplicative FPRAS.

1.2 Related Work

Simple vs. Optimal. We provide an algorithm for the most general setting where an $O(1)$-approximation to the optimal revenue is known using simple mechanisms. It is worth mentioning that a recent result by Dütting et al. [DKL20] shows that simple mechanisms can be used to obtain a $O(\log \log m)$-approximation to the optimal revenue even when the bidders have subadditive valuations. We leave it as an interesting open problem to extend our algorithm to bidders with subadditive valuations.

$(1-\varepsilon)$-Approximation in Item-Independent Settings. We focus on constant factor approximations for general valuations. For more specialized valuations, e.g., unit-demand/additive, there are several interesting results for finding $(1-\varepsilon)$-approximation to the "optimal mechanism". For example, PTASes are known if we restrict our attention to finding the optimal simple mechanism for a single bidder, e.g., item-pricing [CD11] or partition mechanisms [Rub16]. For multiple bidders, PTASes are known for bidders with additive valuations under extra assumptions on distributions (such as i.i.d., MHR,[7] etc.) [DW12, CH13]. The only result that does not require simplicity of the mechanism or extra assumptions on the distribution is [KSM$^+$19], but their algorithm is only a quasi-polynomial time approximation scheme (QPTAS) and computes a $(1-\varepsilon)$-approximation to the optimal revenue for a single unit-demand bidder.

Structured Distributions beyond Item-Independence. When the type distributions can be represented as other structured distributions such as Bayesian networks, Markov Random Fields, or Topic Models, recent results show how to utilize the structure to improve the learnability, approximability, and communication complexity of multi-item auctions [BCD20, CO21, CD21]. We believe that tools developed in this work would be useful to obtain similar improvement in terms of the computational complexity for computing approximately optimal mechanisms for structured distributions beyond item-independence.

2 Preliminaries

We focus on revenue maximization in the combinatorial auction with n independent bidders and m heterogeneous items. We denote bidder i's type t_i as $\{t_{ij}\}_{j\in[m]}$, where t_{ij} is bidder i's private information about item j. For each i, j, we assume t_{ij} is drawn independently from the distribution D_{ij}. Let $D_i = \times_{j=1}^m D_{ij}$ be the distribution of bidder i's type and $D = \times_{i=1}^n D_i$ be the distribution of the type profile. We only consider discrete distributions in this paper. We use $\mathcal{T}_{ij}$ (or $\mathcal{T}_i, \mathcal{T}$) and f_{ij} (or f_i, f) to denote the support and the probability mass function of D_{ij} (or D_i, D). For notational convenience, we let t_{-i} to be the types of all bidders except i and $t_{<i}$ (or $t_{\leq i}$) to be the types of the first $i-1$ (or i) bidders. Similarly, we define D_{-i}, $\mathcal{T}_{-i}$ and f_{-i} for the corresponding distribution, support of the distribution, and probability mass function.

Valuation Functions. For every bidder i, denote her valuation function as $v_i(\cdot,\cdot) : \mathcal{T}_i \times 2^{[m]} \to \mathbb{R}_+$. For every $t_i \in \mathcal{T}_i, S \subseteq 2^{[m]}$, $v_i(t_i, S)$ is bidder i's value for receiving a set S of items, when her type is t_i. In the paper, we are interested in constrained-additive and XOS valuations. For every $i \in [n]$, bidder i's valuation $v_i(\cdot,\cdot)$ is *constrained-additive* if the bidder can receive a set of items subject to some downward-closed feasibility constraint $\mathcal{F}_i$. Formally, $v_i(t_i, S) = \max_{R\in 2^S\cap\mathcal{F}} \sum_{j\in R} t_{ij}$ for every type t_i and set S. It contains classic valuations such as additive ($\mathcal{F}_i = 2^{[m]}$) and unit-demand ($\mathcal{F}_i = \cup_{j\in[m]}\{j\}$). For constrained-additive valuations, we use t_{ij} to denote bidder i's value for item j. For every $i \in [n]$, bidder i's valuation $v_i(\cdot,\cdot)$ is *XOS (or fractionally-subadditive)* if each t_{ij} represents a set of K non-negative numbers $\{\alpha_{ij}^{(k)}(t_{ij})\}_{k\in[K]}$, for some integer K, and $v_i(t_i, S) = \max_{k\in[K]} \sum_{j\in S} \alpha_{ij}^{(k)}(t_{ij})$, for every type t_i and set S. We denote by

[7] That is, $f_{ij}(v)/1 - F_{ij}(v)$ is monotone non-decreasing (MHR) for each i, j, where f_{ij} is the pdf and F_{ij} is the cdf.

$V_{ij}(t_i) = v_i(t_i, \{j\})$ the value for a single item j. Since the value of the bidder for item j only depends on t_{ij}, we denote $V_{ij}(t_{ij})$ as the singleton value.

Mechanisms. A mechanism $\mathcal{M}$ can be described as a tuple (σ, p), where σ is the *interim allocation* rule of $\mathcal{M}$ and p stands for the payment rule. Formally, for every bidder i, type t_i and set S, $\sigma_{iS}(t_i)$ is the interim probability that bidder i with type t_i receives exact bundle S. We use standard concepts of BIC, DSIC and IR for mechanisms. See Appendix A for the formal definitions. For any BIC and IR mechanism $\mathcal{M}$, denote $\text{REV}(\mathcal{M})$ the revenue of $\mathcal{M}$. Denote OPT the optimal revenue among all BIC and IR mechanisms. Throughout this paper, the two classes of simple mechanisms we focus on are *rationed posted price* (RPP) mechanisms and *two-part tariff* (TPT) mechanisms, which are both described in Section 1. We denote PREV the optimum revenue achievable among all RPP mechanisms.

Access to the Bidders' Valuations. We define several ways to access a bidder's valuation.

Definition 1 (Value and Demand Oracle). *A* value oracle *for a valuation $v(\cdot,\cdot)$ takes a type t and a set of items $S \subseteq [m]$ as input, and returns the bidder's value $v(t, S)$ for the bundle S. A* demand oracle *for a valuation $v(\cdot,\cdot)$ takes a type t and a collection of non-negative prices $\{p_j\}_{j\in[m]}$ as input, and returns a utility-maximizing bundle, i.e. $S^* \in \arg\max_{S\subseteq[m]} \left(v(t,S) - \sum_{j\in S} p_j\right)$. In this paper, we use $\text{DEM}_i(\cdot,\cdot)$ to denote the demand oracle for bidder i's valuation $v_i(\cdot,\cdot)$.*

For constrained-additive valuations, our result only requires query access to a value oracle and a *demand oracle* for every bidder i's valuation $v_i(\cdot,\cdot)$. For XOS valuations, we need a stronger demand oracle that allows "scaled types" as input. We refer to the stronger oracle as the *adjustable demand oracle*.

Definition 2 (Adjustable Demand Oracle). *An* adjustable demand oracle *for bidder i's XOS valuation $v_i(\cdot,\cdot)$ takes a type t, a collection of non-negative coefficients $\{b_j\}_{j\in[m]}$, and a collection of non-negative prices $\{p_j\}_{j\in[m]}$ as input. For every item j, b_j is a scaling factor for t_{ij}, meaning that each of the K numbers $\{\alpha_{ij}^{(k)}(t_{ij})\}_{k\in[K]}$, i.e. the contribution of item j under each additive function, is multiplied by b_j. The oracle outputs a favorite bundle S^* with respect to the adjusted contributions and the prices $\{p_j\}_{j\in[m]}$, as well as the additive function $\{\alpha_{ij}^{(k^*)}(t_{ij})\}_{j\in[m]}$ for some $k^* \in [K]$ that achieves the highest value on S^*. Formally, $(S^*, k^*) \in \arg\max_{S\subseteq[m],k\in[K]} \left\{\sum_{j\in S} b_j\alpha_{ij}^{(k)}(t_{ij}) - \sum_{j\in S} p_j\right\}$. We use $\text{ADEM}_i(\cdot,\cdot,\cdot)$ to denote the adjustable demand oracle for bidder i's XOS valuation $v_i(\cdot,\cdot)$.*

The adjustable demand oracle can be viewed as a generalization of the demand oracle for XOS valuations. In the above definition, if every coefficient b_j is 1, then the adjustable demand oracle outputs the utility-maximizing bundle S^* (as in the demand oracle) and the additive function that achieves the value for this set. For general b_j's, the adjustable demand oracle scales item j's contribution to bidder i's value by a b_j factor. The output bundle S^* maximizes the adjusted utility.[8]

Definition 3 (Bit Complexity of an Instance). *Given any instance of our problem represented as the tuple $(\mathcal{T}, D, v = \{v_i(\cdot,\cdot)\}_{i\in[n]})$, Denote as b_f the bit complexity of elements in $\{f_{ij}(t_{ij})\}_{i\in[n],j\in[m],t_{ij}\in\mathcal{T}_{ij}}$. For constrained-additive valuations, denote as b_v the bit complexity of elements in $\{t_{ij}\}_{i\in[n],j\in[m],t_{ij}\in\mathcal{T}_{ij}}$. For XOS valuations, denote as b_v the bit complexity of elements in $\{\alpha_{ij}^{(k)}(t_{ij})\}_{i\in[n],j\in[m],t_{ij}\in\mathcal{T}_{ij},k\in[K]}$. We define the value $\max(b_v, b_f)$ to be the bit complexity of the instance.*

3 Linear Program Relaxation via Lifting

In this section, we present the linear program relaxation for computing an approximately optimal simple mechanism. The main result of our paper is as follows:

[8]Note that for every collection of scaling factors, the query to the adjusted demand oracle is simply a demand query for a different XOS valuation. If all additive functions of t_i are explicitly given, then the adjusted demand oracle can be simulated in time $O(mK)$.

Theorem 1. *Let $T = \sum_{i,j} |\mathcal{T}_{ij}|$ and b be the bit complexity of the problem instance (Definition 3). For any $\delta > 0$, there exists an algorithm that computes a RPP mechanism or a TPT mechanism, such that the revenue of the mechanism is at least $c \cdot \text{OPT}$ for some absolute constant $c > 0$ with probability $1 - \delta - \frac{2}{nm}$. For constrained-additive valuations, our algorithm assumes query access to a value oracle and a demand oracle of bidders' valuations. For XOS valuations, our algorithm assumes query access to a value oracle and an adjustable demand oracle. The algorithm has running time poly$(n, m, T, b, \log(1/\delta))$.*

For any matroid-rank valuation, i.e., the downward-closed feasibility constraint is a matroid, the value and demand oracle can be simulated in polynomial time using greedy algorithms. For more general constraints, it is standard to assume access to the value and demand oracle. We also show that the adjustable demand oracle (rather than a demand oracle) is necessary to obtain our XOS result. In Theorem 14, we prove that (even an approximation of) ADEM_i can not be implemented in polynomial time, given access to the value oracle, demand oracle, and XOS oracle.

As most of the technical barriers already exist in the constrained-additive case, for exposition purposes, we focus on constrained-additive valuations in the main body (unless explicitly stated).[9] Before stating our LP, we first provide a brief recap of the existential result by Cai and Zhao [CZ17] summarized in Lemma 1.[10]

Definition 4. *For any $i \in [n], j \in [m]$, and any feasible[11] interim allocation σ, and non-negative numbers $\tilde{\boldsymbol{\beta}} = \{\tilde{\beta}_{ij} \in \mathcal{T}_{ij}\}_{i\in[n],j\in[m]}$, $\mathbf{c} = \{c_i\}_{i\in[n]}$ and $\mathbf{r} = \{r_{ij}\}_{i\in[n],j\in[m]} \in [0,1]^{nm}$ (referred to as the dual parameters), let $\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}, \mathbf{c}, \mathbf{r})$ be the welfare under allocation σ truncated at $\tilde{\beta}_{ij} + c_i$ for every i, j. Formally,*

$$\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}, \mathbf{c}, \mathbf{r}) = \sum_i \sum_{t_i} f_i(t_i) \cdot \sum_{S\subseteq[m]} \sigma_{iS}(t_i) \sum_{j\in S} t_{ij} \cdot \left(\mathbb{1}[t_{ij} < \tilde{\beta}_{ij} + c_i] + r_{ij} \cdot \mathbb{1}[t_{ij} = \tilde{\beta}_{ij} + c_i]\right).$$

Lemma 1. *[CZ17] Given any BIC and IR mechanism $\mathcal{M}$ with interim allocation σ, where $\sigma_{iS}(t_i)$ is the interim probability for bidder i to receive exactly bundle S when her type is t_i, there exist non-negative numbers $\tilde{\boldsymbol{\beta}}^{(\sigma)} = \{\tilde{\beta}_{ij}^{(\sigma)} \in \mathcal{T}_{ij}\}_{i\in[n],j\in[m]}$, $\mathbf{c}^{(\sigma)} = \{c_i^{(\sigma)}\}_{i\in[n]}$ and $\mathbf{r}^{(\sigma)} \in [0,1]^{nm}$ that satisfy[12]*

1. $\sum_{i\in[n]} \left(\Pr_{t_{ij}}[t_{ij} > \tilde{\beta}_{ij}^{(\sigma)}] + r_{ij}^{(\sigma)} \cdot \Pr_{t_{ij}}[t_{ij} = \tilde{\beta}_{ij}^{(\sigma)}]\right) \le \frac{1}{2}, \forall j,$

2. $\frac{1}{2} \cdot \sum_{t_i\in\mathcal{T}_i} f_i(t_i) \cdot \sum_{S:j\in S} \sigma_{iS}(t_i) \le \Pr_{t_{ij}}[t_{ij} > \tilde{\beta}_{ij}^{(\sigma)}] + r_{ij}^{(\sigma)} \cdot \Pr_{t_{ij}}[t_{ij} = \tilde{\beta}_{ij}^{(\sigma)}], \forall i, j,$

3. $\sum_{i\in[n]} c_i^{(\sigma)} \le 8 \cdot \text{PREV},$

and the corresponding $\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)}, \mathbf{r}^{(\sigma)})$ satisfies the following inequalities:

4. $\text{REV}(\mathcal{M}) \le 28 \cdot \text{PREV} + 4 \cdot \text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)}, \mathbf{r}^{(\sigma)}),$

5. $\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)}, \mathbf{r}^{(\sigma)}) \le 64 \cdot \text{PREV} + 8 \cdot \text{REV}(\mathcal{M}_1^{(\sigma)})$, *where $\mathcal{M}_1^{(\sigma)}$ is some TPT mechanism.*

Remark 1. *For continuous type distributions, there exists $\tilde{\boldsymbol{\beta}}^{(\sigma)}$ that satisfy both Property 1 and 2 of Lemma 1 with $r_{ij}^{(\sigma)} = 1, \forall i, j$ for every σ. For discrete distributions, such a $\tilde{\boldsymbol{\beta}}^{(\sigma)}$ may not exist. This is simply a tie-breaking issue, and the role of $\mathbf{r}^{(\sigma)}$ is to fix it. Roughly speaking, $r_{ij}^{(\sigma)}$ is the probability that bidder i wins item j, when she is indifferent between purchasing or not. Readers can treat $\mathbf{r}^{(\sigma)}$ as the all-one vector to get the intuition behind our approach.*

[9] The linear program for XOS valuations can be found in Figure 4 in Appendix C.3.

[10] The statement is for constrained-additive bidders. See Appendix C.1 for the statement for XOS bidders.

[11] For constrained-additive bidders, an interim allocation σ is feasible if it can be implemented by a mechanism whose allocation rule always respects all bidders' feasibility constraints. It is without loss of generality to consider feasible interim allocations.

[12] [CZ17] provides an explicit way to calculate $\tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)}, \mathbf{r}^{(\sigma)}$. We only include the crucial properties of these parameters here.

By combining Property 4 and 5 of Lemma 1, Cai and Zhao [CZ17] proved that the revenue of any BIC, IR mechanism $\mathcal{M}$ is bounded by a constant number of PREV and the revenue of some TPT mechanism. Recall that PREV is the optimal revenue achieved by an RPP mechanism, which is exactly the Sequential Posted Price mechanism if we restrict the bidders' valuations to unit-demand. Thus we can compute a set of posted prices that approximates PREV by Chawla et al. [CHMS10a].

3.1 Tour to Our Relaxation

To facilitate our discussion about the key components and the intuition behind the relaxation, we present the development of our relaxation and along the way examine several failed attempts. In Theorem 3, we show that the optimal solution of the relaxed problem can indeed be approximated by simple mechanisms. Due to space limitations, we do not include details on the approximation analysis in this section, but focus on our intuition behind each step of our relaxation. Interested readers can find the proof of Theorem 3 in Appendix C.6. We also assume r_{ij} to be 1 for every i and j to keep the notation light.

Step 0: Replace Revenue with the Duality-Based Benchmark. Instead of optimizing the revenue, we optimize the upper bound of revenue. As guaranteed by Lemma 1, for any BIC and IR mechanism $\mathcal{M} = (\sigma, p)$, its revenue is upper bounded by $O(\text{PREV} + \text{CORE}(\sigma, \theta(\sigma)))$, where we use $\theta(\sigma)$ to denote the set of dual parameters $(\tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)})$ guaranteed to exist by Lemma 1. Since we can approximate PREV, it suffices to first approximately maximize $\text{CORE}(\sigma, \theta(\sigma))$ over all feasible interim allocations σ, then compute the TPT in Lemma 1 based on the computed σ. $\text{CORE}(\sigma, \theta(\sigma))$ is the truncated welfare, but the truncation depends on σ in a complex way, causing the function to be highly non-concave in σ (Example 1).

Step 1: Two-Stage Optimization. To overcome the barrier mentioned above, we consider a two-stage optimization problem (Figure 1) by *switching the order of dependence* between the interim allocation σ and dual parameters $\theta = (\boldsymbol{\beta}, \mathbf{c})$. In Stage I, we optimize some function H over the dual parameters $\theta = (\boldsymbol{\beta}, \mathbf{c})$, where $H(\boldsymbol{\beta}, \mathbf{c})$ is the optimum of the Stage II problem for every fixed set of parameters $(\boldsymbol{\beta}, \mathbf{c})$. Constraint (1) and (2) in the Stage I problem are due to Property 1 and 3 of Lemma 1 respectively. In Stage II, for any fixed set of parameters $\theta = (\boldsymbol{\beta}, \mathbf{c})$, we optimize $\text{CORE}(\sigma, \theta)$ over all feasible σ such that the tuple $(\sigma, \boldsymbol{\beta}, \mathbf{c})$ satisfy Property 1, 2, and 3 of Lemma 1. We choose the interim allocation σ as the variables, $\text{CORE}(\sigma, \boldsymbol{\beta}, \mathbf{c})$ as the objective, and include Constraint (4), which corresponds to Property 2 of Lemma 1. Why is the two-stage optimization a relaxation? For any interim allocation σ, (i) the corresponding set of dual parameters $\theta(\sigma)$ is a feasible solution of the first-stage optimization problem, and (ii) σ is feasible in the second-stage optimization w.r.t. $\theta(\sigma)$, so $(\theta(\sigma), \sigma)$ is a feasible solution of the two-stage optimization problem.

Stage I:

$$\textbf{max } H(\boldsymbol{\beta}, \mathbf{c})$$

$$\textbf{s.t.} \quad (1) \quad \sum_{i\in[n]} \Pr_{t_{ij}}[t_{ij} \geq \beta_{ij}] \leq \frac{1}{2} \qquad \forall j$$

$$(2) \quad \sum_{i\in[n]} c_i \leq 8 \cdot \text{PREV}$$

Stage II:

$$H(\boldsymbol{\beta}, \mathbf{c}) = \textbf{max} \sum_{i\in[n]} \sum_{t_i\in\mathcal{T}_i} f_i(t_i) \cdot \sum_{S\subseteq[m]} \sigma_{iS}(t_i) \sum_{j\in S} t_{ij} \cdot \mathbb{1}[t_{ij} \leq \beta_{ij} + c_i]$$

$$\textbf{s.t.} \quad (3) \quad \sigma \text{ is feasible}$$

$$(4) \quad \frac{1}{2} \sum_{t_i\in\mathcal{T}_i} f_i(t_i) \cdot \sum_{S:j\in S} \sigma_{iS}(t_i) \leq \Pr_{t_{ij}}[t_{ij} \geq \beta_{ij}] \qquad \forall i, j$$

Figure 1: Two-stage Optimization over $\theta = (\boldsymbol{\beta}, \mathbf{c})$ and the allocation σ

We now focus on the Stage II problem and try to solve it efficiently for a fixed set of parameters θ. The objective is a linear function of the variables σ, yet the set of variables $\sigma = \{\sigma_{iS}(t_i)\}_{i\in[n], S\subseteq[m], t_i\in\mathcal{T}_i}$ has exponential size. Luckily, the problem can be expressed more succinctly. For any interim allocation σ and dual parameters $\theta = (\boldsymbol{\beta}, \mathbf{c})$, the objective ($\textsc{Core}(\sigma,\theta)$) can be simplified as follows:

$$\textsc{Core}(\sigma,\theta) = \sum_{\substack{i\in[n]\\ t_i\in\mathcal{T}_i}} f_i(t_i) \sum_{S\subseteq[m]} \sigma_{iS}(t_i) \sum_{j\in S} t_{ij}\cdot \mathbb{1}[t_{ij}\le \beta_{ij}+c_i] \;=\; \sum_{\substack{i\in[n],j\in[m]\\ t_{ij}\in\mathcal{T}_{ij}}} \widehat{w}_{ij}(t_{ij})t_{ij}\cdot\mathbb{1}[t_{ij}\le\beta_{ij}+c_i], \tag{1}$$

where $\widehat{w}_{ij}(t_{ij}) = f_{ij}(t_{ij})\cdot\sum_{t_{i,-j}} f_{i,-j}(t_{i,-j})\cdot\sum_{S:j\in S}\sigma_{iS}(t_{ij},t_{i,-j})$ for every $i\in[n], j\in[m], t_{ij}\in\mathcal{T}_{ij}$. We refer to $\{\widehat{w}_{ij}(t_{ij})\}_{i\in[n],j\in[m],t_{ij}\in\mathcal{T}_{ij}}$ as the **marginal reduced form** of the interim allocation rule σ. $\widehat{w}_{ij}(t_{ij})$ represents the probability that bidder i's value for item j is t_{ij} and she receives item j, and the probability is taken over the randomness of the allocation, other bidders' types, as well as her own values for all the other items. Now for every fixed dual parameters θ, $\textsc{Core}$ is expressed as a linear function of the much more succinct representation $\widehat{w} = \{\widehat{w}_{ij}(t_{ij})\}_{i,j,t_{ij}}$ that has polynomial description size. We rewrite the Stage II problem as an LP using the variables $\widehat{w}$. Denote $\textsc{Core}(\hat{w},\theta)$ the last term of Equation (1), which is the objective of the problem. By the definition of $\widehat{w}$, Constraint (4) is equivalent to

$$\frac{1}{2}\cdot\sum_{t_{ij}\in\mathcal{T}_{ij}}\widehat{w}_{ij}(t_{ij}) \le \Pr_{t_{ij}}[t_{ij}\ge\beta_{ij}], \qquad \forall i,j \tag{2}$$

which is a linear constraint on $\widehat{w}$. Let $\mathcal{P}_1$ be the convex polytope that contains all marginal reduced forms $\widehat{w}$ that can be implemented by some feasible allocation σ (corresponds to Constraint (3)) and $\mathcal{P}_2$ be the set of all $\widehat{w}$ that satisfy all constraints in Equation (2). The Stage II problem is equivalent to the LP $\max_{\widehat{w}\in\mathcal{P}_1\cap\mathcal{P}_2}\textsc{Core}(\widehat{w},\theta)$. Unfortunately, since $\mathcal{P}_1$ does not have an explicit succinct description or an efficient separation oracle, it is unclear if the problem can be solved efficiently.

Step 2: Marginal Reduced Form Relaxation. To overcome this barrier, we consider a relaxation of $\mathcal{P}_1$, where the feasibility constraint is only enforced on each bidder separately. We refer to this step as the *marginal reduced form relaxation*. We use $\widehat{w}_i = \{\widehat{w}_{ij}(t_{ij})\}_{j\in[m],t_{ij}\in\mathcal{T}_{ij}}$ to denote a feasible *single-bidder marginal reduced form* for bidder i. Formally, we define the feasible region W_i of $\widehat{w}_i$ in Definition 5.

Definition 5 (Constrained-additive valuations: single-bidder marginal reduced form polytope)**.** *For every $i\in[n]$, suppose bidder i has a constrained-additive valuation with feasibility constraint $\mathcal{F}_i$. Bidder i's single-bidder marginal reduced form polytope $W_i\subseteq[0,1]^{\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ is defined as follows: $\widehat{w}_i\in W_i$ if and only if there exists an allocation rule $\{\sigma_S(t_i)\}_{t_i\in\mathcal{T}_i,S\in\mathcal{F}_i}$, i.e., $\sigma_S(t_i)$ is the probability that i receives set S when her type is t_i, such that* **(i)** *$\sum_{S\in\mathcal{F}_i}\sigma_S(t_i)\le 1, \forall t_i\in\mathcal{T}_i$, and* **(ii)** *$\widehat{w}_{ij}(t_{ij}) = f_{ij}(t_{ij})\cdot\sum_{t_{i,-j}} f_{i,-j}(t_{i,-j})\cdot\sum_{S:j\in S}\sigma_S(t_i)$, for all $j\in[m]$ and $t_{ij}\in\mathcal{T}_{ij}$.*

Throughout this section, we assume access to a separation oracle of W_i for every bidder i. In Theorem 2, we present an efficient separation oracle for another polytope $\widehat{W_i}$ that is a multiplicative approximation to W_i, i.e., $\widehat{W_i}$ is sandwiched between $c\cdot W_i$ and W_i for some absolute constant $c\in(0,1)$, using only queries to bidder i's demand oracle. We will argue later that we can efficiently approximate our problem with the separation oracle for $\widehat{W_i}$.

Here is our relaxation to the (rewritten) Stage II problem: Instead of forcing $\widehat{w}$ to be *implementable jointly* ($\widehat{w}\in\mathcal{P}_1$), we consider the relaxed region $\mathcal{P}'\supseteq\mathcal{P}_1$: $\widehat{w}\in\mathcal{P}'$ if and only if: (i) $\widehat{w}_i\in W_i$, for all bidder $i\in[n]$, and (ii) $\sum_i\sum_{t_{ij}}\widehat{w}_{ij}(t_{ij})\le 1, \forall j\in[m]$. In other words, $\mathcal{P}'$ guarantees that, for every bidder i, $\widehat{w}_i$ is a feasible single-bidder marginal reduced form for i, and the supply constraint is met in terms of marginal reduced forms (rather than ex-post allocations).

The main benefit of this relaxation is computational. Without the relaxation, we need a multiplicative approximation of $\mathcal{P}_1$. Theorem 2 provides such an approximation if we can exactly maximizes the social welfare – a computational task that is substantially harder than answering demand queries. Indeed, we are not aware of any efficient algorithm that exactly maximizes the social welfare with only access to demand oracles of every bidder. The relaxed problem $\max_{\widehat{w}\in\mathcal{P}'\cap\mathcal{P}_2}\textsc{Core}(\widehat{w},\theta)$ is captured by the LP in Figure 2.[13]

[13] We omit the supply constraint $\sum_i\sum_{t_{ij}}\widehat{w}_{ij}(t_{ij})\le 1$ as it is implied by Constraint (1) in the Stage I problem and Constraint (4).

Relaxed Stage II:

$$H(\boldsymbol{\beta}, \mathbf{c}) = \mathbf{max} \sum_{i\in[n]} \sum_{j\in[m]} \sum_{t_{ij}\in\mathcal{T}_{ij}} \widehat{w}_{ij}(t_{ij}) \cdot t_{ij} \cdot \mathbb{1}[t_{ij} \le \beta_{ij} + c_i]$$

$$\textbf{s.t.} \quad (3) \quad \widehat{w}_i \in W_i \qquad \forall i$$

$$(4) \quad \frac{1}{2}\sum_{t_{ij}} \widehat{w}_{ij}(t_{ij}) \le \Pr_{t_{ij}}[t_{ij} \ge \beta_{ij}] \qquad \forall i, j$$

$$\widehat{w}_{ij}(t_{ij}) \ge 0 \qquad \forall i, j, t_{ij}$$

Figure 2: The Relaxed Stage II Problem over the Marginal Reduced Forms

Consider the two-stage optimization with the relaxed Stage II problem. For every fixed parameters θ, the relaxed Stage II problem can be solved efficiently (assuming a separation oracle of W_i for every i). Unfortunately, we do not know how to solve the two-stage optimization problem efficiently, as the number of different dual parameters is exponential in n and m, and enumerating through all possible choices of dual parameters is not an option. To overcome this obstacle, we need ideas explained in the following step.

Step 3: Lifting the problem to a higher dimensional space. Instead of enumerating all possible dual parameters θ, we optimize over *distributions of the parameters*. To guarantee that the number of decision variables in our program is polynomial, we focus on *product distributions* over the parameters. Formally, for every i, j, let $\mathcal{C}_{ij}$ be a distribution over $\mathcal{V}_{ij} \times \Delta$, where $\mathcal{V}_{ij}$ and Δ are the set of possible values of β_{ij} and c_i accordingly, after discretization (See Footnote *a* in Figure 3 for a formal definition). All $\mathcal{C}_{ij}$'s are independent. In our program, we use decision variables $\{\hat{\lambda}_{ij}(\beta_{ij}, \delta_{ij})\}_{i\in[n], j\in[m], \beta_{ij}\in\mathcal{V}_{ij}, \delta_{ij}\in\Delta}$ to represent the distribution $\mathcal{C}_{ij}$, i.e., $\hat{\lambda}_{ij}(a, b) = \Pr_{(\beta_{ij}, \delta_{ij})\sim\mathcal{C}_{ij}}[\beta_{ij} = a \wedge \delta_{ij} = b]$. Notice that if the parameters are drawn from a product distribution, the parameter "c_i" may be different for each item j. To distinguish them, we use δ_{ij} to replace the original parameter c_i in our program.

Now we maximize the expected value of the CORE function over all product distributions $\times_{i,j}\mathcal{C}_{ij}$ (represented by decision variables $\hat{\lambda}$) and the allocations (represented by the marginal reduced form $\widehat{w}$). If the parameters θ and allocation $\widehat{w}$ are generated independently, the expected CORE is not a linear objective, since the contributed truncated welfare in CORE is $\widehat{w}_{ij}(t_{ij}) \cdot \hat{\lambda}_{ij}(\beta_{ij}, \delta_{ij}) \cdot t_{ij} \cdot \mathbb{1}[t_{ij} \le \beta_{ij} + \delta_{ij}]$ for every $t_{ij}, \beta_{ij}, \delta_{ij}$. *To linearize the objective, we lift the problem to a higher dimensional space and consider joint distributions over the parameters and allocations.* We do not consider arbitrary joint distributions, and only focus on the ones that correspond to the following generative process: first draw (β, δ) from a product distribution (according to $\hat{\lambda}$), then choose a feasible allocation $\widehat{w}^{(\beta,\delta)} = \{\widehat{w}_{ij}^{(\beta,\delta)}(t_{ij})\}_{i,j,t_{ij}}$ conditioned on (β, δ). Since there are too many parameters (β, δ), we certainly cannot afford to store all $\widehat{w}^{(\beta,\delta)}$'s explicitly. Instead, for each bidder i and item j we introduce a new set of decision variables $\{\lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij})\}_{t_{ij}\in\mathcal{T}_{ij}, \beta_{ij}\in\mathcal{V}_{ij}, \delta_{ij}\in\Delta}$, where $\lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij})$ is the marginal probability for the following three events to happen simultaneously in our generative process: **(a)** $(\beta_{ij}, \delta_{ij})$ are the parameters for i and j. **(b)** Bidder i receives item j. **(c)** Bidder i's value for item j is t_{ij}. Formally,

$$\lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij}) = \hat{\lambda}_{ij}(\beta_{ij}, \delta_{ij}) \cdot \sum_{\{(\beta_{i'j'}, \delta_{i'j'})\}_{(i',j')\neq(i,j)}} \left(\widehat{w}_{ij}^{(\beta,\delta)}(t_{ij})/f_{ij}(t_{ij})\right) \cdot \prod_{(i',j')\neq(i,j)} \hat{\lambda}_{i'j'}(\beta_{i'j'}, \delta_{i'j'}) \quad (3)$$

With the new variables $\lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij})$'s, we can express the objective as an linear function:

$$\sum_{i\in[n]} \sum_{j\in[m]} \sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij}) \cdot t_{ij} \cdot \sum_{\beta_{ij}\in\mathcal{V}_{ij}, \delta_{ij}\in\Delta} \lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij}) \cdot \mathbb{1}[t_{ij} \le \beta_{ij} + \delta_{ij}].$$

Our program can be viewed as an "expected version" of the the two-stage optimization, when the parameters $\theta = (\beta, \delta) \sim \times_{i,j} \mathcal{C}_{ij}$. In other words, we only require the constraints to be satisfied in expectation. We discuss our relaxation in more details in Section 3.2.

3.2 Our LP and a Sketch of the Proof of Theorem 1

We present a sketch of the proof of Theorem 1 for constrained-additive bidders and our main linear program (Figure 3). Although the LP has many constraints and may seem intimidating at first, all constraints follow quite naturally from our derivation in Section 3.1. See Section 3.3 for more details.

The first step of our proof is to estimate PREV using Lemma 2 from [CHMS10a].

Lemma 2 (Theorem 14 and Appendix F in [CHMS10a]). *There exists an algorithm that with probability at least $1-\frac{2}{nm}$, computes a Rationed Posted Price mechanism $\mathcal{M}$ such that* $\text{REV}(\mathcal{M}) \geq \frac{1}{6.75}(1-\frac{1}{nm}) \cdot \text{PREV}$. *The algorithm runs in time poly*$(n, m, \sum_{i,j} |\mathcal{T}_{ij}|)$.

Denote $\mathcal{E}$ the event that an RPP in Lemma 2 is computed successfully. For simplicity, we will condition on the event $\mathcal{E}$ for the rest of this section. Let $\widetilde{\text{PREV}}$ be the revenue of the RPP mechanism found in Lemma 2.

Next, we argue that the LP in Figure 3 (or Figure 4 when the valuations are XOS) can be solved efficiently. Note that there are poly$(n, m, \sum_{i,j} |\mathcal{T}_{ij}|)$ constraints except for Constraint (1), where we need to enforce the feasibility of single-bidder marginal reduced forms. It suffices to construct an efficient separation oracle for W_i for every i. However, to the best of our knowledge, W_i does not have a succinct explicit description or an efficient separation oracle. For constrained-additive valuations, we construct another polytope $\widehat{W}_i$ such that: (i) $\widehat{W}_i$ is a multiplicative approximation of W_i, i.e., $c \cdot W_i \subseteq \widehat{W}_i \subseteq W_i$ for some absolute constant $c \in (0,1)$, and (ii) There exists an efficient separation oracle for $\widehat{W}_i$ given access to the demand oracle.

Theorem 2. *Let $T = \sum_{i,j} |\mathcal{T}_{ij}|$ and b be the bit complexity of the problem instance (Definition 3). For any $i \in [n]$ and $\delta \in (0,1)$, there is an algorithm that constructs a convex polytope $\widehat{W}_i \in [0,1]^{\sum_{j\in[m]} |\mathcal{T}_{ij}|}$ using poly$(n, m, T, \log(1/\delta))$ samples from D_i, such that with probability at least $1-\delta$,*

1. *$\frac{1}{12} \cdot W_i \subseteq \widehat{W}_i \subseteq W_i$, and the vertex-complexity (Definition 10) of $\widehat{W}_i$ is poly$(n, m, T, b, \log(1/\delta))$.*
2. *There exists a separation oracle SO for $\widehat{W}_i$, given access to the demand oracle for bidder i's valuation. The running time of SO on any input with bit complexity b' is poly$(n, m, T, b, b', \log(1/\delta))$ and makes poly$(n, m, T, b, b', \log(1/\delta))$ queries to the demand oracle.*

The algorithm constructs the polytope and the separation oracle SO in time poly$(n, m, T, b, \log(1/\delta))$.

Indeed, we prove a more general result regarding polytopes that can be expressed as a "mixture of polytopes" (Theorem 8), which can be viewed as a generalization of the technique developed in [CDW12c] for approximating the polytope of all feasible reduced forms. We postpone the proof of Theorem 2 to Appendix D.2.

To solve the LP relaxation, we replace W_i by $\widehat{W}_i$ in the LP in Figure 3 for every $i \in [n]$, and solve the LP in polynomial time using the ellipsoid method. Clearly, this solution is also feasible for the original LP in Figure 3. Moreover, since $\widehat{W}_i$ contains $c \cdot W_i$, we can show that **the objective value of the solution we computed is at least $c \cdot \textbf{OPT}_{\textbf{LP}}$**, where OPT_{LP} the optimum of the LP in Figure 3. Our proof of Theorem 2 heavily relies on the fact that W_i is a down-monotone polytope,[14] which does not hold in the XOS case. For XOS valuations, we construct the polytope $\widehat{W}_i$ with a weaker guarantee: For every vector x in W_i, there exists another vector x' in $\widehat{W}_i$ such that for every coordinate j, $x_j/x'_j \in [a,b]$ for some absolute constant $0 < a < b$, and vice versa. See Appendix E for a complete proof of Theorem 1 (including the XOS case).

Next, we argue that the LP optimum can be approximated by simple mechanisms. [CZ17] shows that for any BIC and IR mechanism $\mathcal{M}$, $\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)}, \mathbf{r}^{(\sigma)})$ (as stated in Lemma 1) can be bounded by a constant number of PREV and the revenue of a TPT (see Property 5 of Lemma 1). We generalize their result by proving that for any feasible solution of the LP, its objective can be bounded by (a constant times) the revenue of a RPP or TPT mechanism, and the mechanism can be computed efficiently given the feasible solution.

Definition 6. *Let $(w, \lambda, \hat{\lambda}, \boldsymbol{d} = (d_i)_{i\in[n]})$ be any feasible solution of the LP in Figure 3. For every $j \in [m]$, define* $Q_j = \frac{1}{2} \cdot \sum_{i\in[n]} \sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij}) \cdot t_{ij} \cdot \sum_{\beta_{ij}\in\mathcal{V}_{ij}, \delta_{ij}\in\Delta} \lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij}) \cdot \mathbb{1}[t_{ij} \leq \beta_{ij} + \delta_{ij}]$.[15]

[14] A polytope $\mathcal{P} \subseteq [0,1]^d$ is down-monotone if and only if for every $\boldsymbol{x} \in \mathcal{P}$ and $\mathbf{0} \leq \boldsymbol{x}' \leq \boldsymbol{x}$, we have $\boldsymbol{x}' \in \mathcal{P}$.

[15] Recall that $\lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij})$ is introduced in Step 3 of Section 3.1. See Figure 3 for the formal definition.

Clearly, for any feasible solution of the LP, the objective function is $2\cdot\sum_{j\in[m]}Q_j$. We prove in Theorem 3 that $2\cdot\sum_{j\in[m]}Q_j$ can be bounded by the revenue of $\mathcal{M}_{\text{TPT}}$ (Mechanism 1) and the RPP $\mathcal{M}_{\text{PP}}$ (Lemma 2). As we can efficiently compute a feasible solution whose objective is $\Omega(\text{OPT}_{\text{LP}})$, Theorem 3 implies that we can compute in polynomial time a simple mechanism whose revenue is at least $\Omega(\text{OPT}_{\text{LP}}+\text{PREV})$.

Theorem 3. *Let $(w,\lambda,\hat{\lambda},\boldsymbol{d})$ be any feasible solution of the LP in Figure 3. Let $\mathcal{M}_{PP}$ be the rationed posted price mechanism computed in Lemma 2. Let $\mathcal{M}_{TPT}$ be the two-part tariff mechanism shown in Mechanism 1 with prices $\{Q_j\}_{j\in[m]}$. Then the objective function of the solution $2\cdot\sum_{j\in[m]}Q_j$ is bounded by $c_1\cdot\text{REV}(\mathcal{M}_{PP})+c_2\cdot\text{REV}(\mathcal{M}_{TPT})$, for some absolute constant $c_1,c_2>0$. Moreover, both $\mathcal{M}_{PP}$ and $\mathcal{M}_{TPT}$ can be computed in time poly$(n,m,\sum_{i,j}|\mathcal{T}_{ij}|)$ with access to the demand oracle for the bidders' valuations.*

The proof of Theorem 3 combines the "shifted CORE" technique by Cai and Zhao [CZ17] with several novel ideas to handle the new challenges due to the relaxation. We postpone it to Appendix C.6.

Mechanism 1 Two-part Tariff Mechanism $\mathcal{M}_{\text{TPT}}$

0: Before the mechanism starts, the seller computes the *price* Q_j (Definition 6) for every item j.
1: Bidders arrive sequentially in the lexicographical order.
2: When bidder i arrive, the seller shows her the set of available items $S_i(t_{<i})\subseteq[m]$, as well as their prices. Note that $S_i(t_{<i})$ is the set of items that are not purchased by the first $i-1$ bidders, which depends on $t_{<i}$. We use $S_1(t_{<1})$ to denote $[m]$.
3: Bidder i is asked to pay an *entry fee*. The seller samples a type $t_i'\sim D_i$, and sets the entry fee as: $\xi_i(S_i(t_{<i}),t_i')=\max_{S'\subseteq S_i(t_{<i})}\left(v_i(t_i',S')-\sum_{j\in S'}Q_j\right)$. The entry fee is bidder i's utility for her favorite set under prices Q_j's if her type is t_i'.
4: If bidder i (with type t_i) agrees to pay the entry fee $\xi_i(S_i(t_{<i}),t_i')$, then she can enter the mechanism and take her favorite set $S^*\in\arg\max_{S'\subseteq S_i(t_{<i})}\left(v_i(t_i,S')-\sum_{j\in S'}Q_j\right)$, by paying $\sum_{j\in S^*}Q_j$. Otherwise, the bidder gets nothing and pays 0.

Remark 2. *Theorem 3 indeed holds even if the bidders arrive in an arbitrary order in $\mathcal{M}_{TPT}$. We choose the lexicographical order only to keep the notation light.*

We complete the last step of our proof by showing $\text{OPT}=O(\text{OPT}_{\text{LP}}+\text{PREV})$ in Lemma 3. More specifically, we show that for any mechanism $\mathcal{M}=(\sigma,p)$, the tuple $(\sigma,\tilde{\boldsymbol{\beta}}^{(\sigma)},\mathbf{c}^{(\sigma)},r^{(\sigma)})$ stated in Lemma 1 corresponds to a feasible solution of the LP in Figure 3 whose objective is at least $\text{CORE}(\sigma,\tilde{\boldsymbol{\beta}}^{(\sigma)},\mathbf{c}^{(\sigma)},r^{(\sigma)})$. Hence, the revenue of $\mathcal{M}$ is upper bounded by PREV and OPT_{LP}. The proof is postponed to Appendix C.4. Indeed, for both Theorem 3 and Lemma 3, we prove a general statement that applies to XOS valuations, which requires a generalized LP and definitions. See Theorem 6 and Lemma 5 in Appendix C for details.

Lemma 3. *For any BIC and IR mechanism $\mathcal{M}$, $\text{REV}(\mathcal{M})\leq 28\cdot\text{PREV}+4\cdot\text{OPT}_{\text{LP}}$.*

3.3 Interpretation of Our LP in Figure 3.

We explain our LP in this section. The objective is the expected CORE, as explained in Step 3 in Section 3.1. According to our definition of $\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})$ and Constraint (3), $\{w_{ij}(t_{ij})\}_{i,j,t_{ij}}$ corresponds to the expected marginal reduced form, that is, $w_{ij}(t_{ij})$ is the expected probability for bidder i to receive item j and her value for item j is t_{ij}. Constraints (1) and (2) simply sets the feasible region of the expected marginal reduced form w. They follow directly from the fact that every realized $\widehat{w}^{(\beta,\delta)}$ is feasible (see Step 3 in Section 3.1). Constraint (4) follows from Equation (3) and the fact that every $\widehat{w}_{ij}^{(\beta,\delta)}(t_{ij})/f_{ij}(t_{ij})$ is in $[0,1]$. Constraint (5) implies that $\{\hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\}_{\beta_{ij},\delta_{ij}}$ correspond to a distribution $\mathcal{C}_{ij}$.

$$\textbf{max} \sum_{i\in[n]}\sum_{j\in[m]}\sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij})\cdot t_{ij}\cdot \sum_{\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}\in\Delta} \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\cdot \mathbb{1}[t_{ij}\leq \beta_{ij}+\delta_{ij}]$$

s.t.

Allocation Feasibility Constraints:

$$(1)\quad w_i \in W_i \qquad \forall i$$

$$(2)\quad \sum_i \sum_{t_{ij}\in\mathcal{T}_{ij}} w_{ij}(t_{ij}) \leq 1 \qquad \forall j$$

Natural Feasibility Constraints:

$$(3)\quad f_{ij}(t_{ij})\cdot \sum_{\beta_{ij}\in\mathcal{V}_{ij}}\sum_{\delta_{ij}\in\Delta} \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij}) = w_{ij}(t_{ij}) \qquad \forall i,j,t_{ij}\in\mathcal{T}_{ij}$$

$$(4)\quad \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij}) \leq \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij}) \qquad \forall i,j,t_{ij},\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}$$

$$(5)\quad \sum_{\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}\in\Delta} \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij}) = 1 \qquad \forall i,j$$

Problem Specific Constraints:

$$(6)\quad \sum_{i\in[n]}\sum_{\beta_{ij}\in\mathcal{V}_{ij}}\sum_{\delta_{ij}\in\Delta} \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\cdot \Pr_{t_{ij}\sim D_{ij}}[t_{ij}\geq\beta_{ij}] \leq \frac{1}{2} \qquad \forall j$$

$$(7)\quad \frac{1}{2}\sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij})\left(\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij}) + \lambda_{ij}(t_{ij},\beta_{ij}^+,\delta_{ij})\right) \leq$$
$$\hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\cdot \Pr_{t_{ij}}[t_{ij}\geq\beta_{ij}] + \hat{\lambda}_{ij}(\beta_{ij}^+,\delta_{ij})\cdot \Pr_{t_{ij}}[t_{ij}\geq\beta_{ij}^+] \qquad \forall i,j,\beta_{ij}\in\mathcal{V}_{ij}^0,\delta_{ij}\in\Delta$$

$$(8)\quad \sum_{\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}\in\Delta} \delta_{ij}\cdot\hat{\lambda}_{ij}(\beta_{ij},\delta_{ij}) \leq d_i \qquad \forall i,j$$

$$(9)\quad \sum_{i\in[n]} d_i \leq 111\cdot\widetilde{\text{PREV}}$$

$$\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\geq 0, \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\geq 0, w_{ij}(t_{ij})\geq 0, d_i\geq 0 \qquad \forall i,j,t_{ij},\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}$$

Variables: [a]

- $\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})$, for all i,j and $t_{ij}\in\mathcal{T}_{ij}$, $\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}\in\Delta$. See Step 3 of Section 3.1 for an explanation of this variable.

- $\hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})$, for all i,j, $\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}\in\Delta$, denoting the distribution $\mathcal{C}_{ij}$ over $\mathcal{V}_{ij}\times\Delta$.

- $w_{ij}(t_{ij})$, for all $i\in[n], j\in[m], t_{ij}\in\mathcal{T}_{ij}$, denoting the expected marginal reduced form. We denote $w_i = \{w_{ij}(t_{ij})\}_{j\in[m],t_{ij}\in\mathcal{T}_{ij}}$ the vector of all variables associated with bidder i.

- d_i, for all $i\in[n]$, denoting an upper bound of the expectation of δ_{ij} over distribution $\mathcal{C}_{ij}$ for all j.

[a] For every i,j, let $\mathcal{V}_{ij}^0=\mathcal{T}_{ij}$ be the set of all possible values of t_{ij}. To address the tie-breaking issue in Remark 1, let $\varepsilon_r>0$ be an arbitrarily small number,and define $\mathcal{V}_{ij}^+=\{t_{ij}+\varepsilon_r : t_{ij}\in\mathcal{T}_{ij}\}$ and $\mathcal{V}_{ij}=\mathcal{V}_{ij}^0\cup\mathcal{V}_{ij}^+$. Let Δ be a geometric discretization of range $[\widetilde{\text{PREV}}/n, 55\cdot\widetilde{\text{PREV}}]$. Formally, $\delta\in\Delta$ if and only if $\delta=\frac{2^x}{n}\cdot\widetilde{\text{PREV}}$ for some integer x such that $0\leq x\leq\lceil\log(55n)\rceil$. Finally, for each $\beta\in\mathcal{V}_{ij}^0$, let $\beta^+=\beta+\varepsilon_r\in\mathcal{V}_{ij}^+$. Note that the LP (or the LP in Figure 4) do not depend on the choice of ε_r, so we can choose ε_r to be sufficiently small. In fact, let b be an upper bound of the bit complexity of the problem instance, and the bit complexity of any feasible solution of our LP. Our proof works as long as $\varepsilon_r<\min\{\frac{1}{2^{\text{poly}(b)}},\frac{\text{PREV}}{\sum_{i,j}|\mathcal{T}_{ij}|}\}$.

Figure 3: LP Relaxation for Constrained-Additive Bidders

Constraints (6) - (9) are specialized for our problem, which guarantees that the LP optimum can be bounded by simple mechanisms. Constraint (6) follows from taking expectation on both sides of Constraint (1) in Figure 1, over the randomness of β_{ij}. Constraints (8) and (9) correspond to Constraint (2) in Figure 1. Here we bound the expectation of δ_{ij} by a unified upper bound d_i for every j.[16] It is worth emphasizing Constraint (7), which corresponds to Constraint (4) in Figure 1 (and Figure 2). Instead of taking expectations over the dual parameters, we force the constraint to hold for every β_{ij} and δ_{ij}. This is an important property that is crucial in our analysis (see Footnote 22 in Lemma 9). Readers may notice that Constraint (2) is implied by Constraints (3), (6) and (7). We keep Constraint (2) so that it is clear that the supply constraint is enforced over the (expected) marginal reduced form. The Problem Specific Constraints (6)-(9) in the LP in Figure 3 are expected versions of the constraints in the two-stage optimization problem in Figure 1, which are directly inspired by the Properties 1, 2, and 3 in Lemma 1. They are crucial to guarantee that optimal value of the LP in Figure 3 is still approximable by simple mechanisms.

4 Sample Access to the Distribution

In this section, we focus on the case where we have only access to the bidders' distribution. Our goal is again to compute an approximately optimal mechanism. Our plan is as follows: (i) for each $i \in [n]$ and $j \in [m]$, take $O(\log(1/\delta)/\varepsilon^2)$ samples from D_{ij}, and let $\widehat{D}$ be the uniform distribution over the samples. By the DKW inequality [DKW56], $\widehat{D}_{ij}$ and D_{ij} have Kolmogorov distance (Definition 30) no more than ε with probability at least $1-\delta$. (ii) We then apply our algorithm in Theorem 1 to compute an RPP or TPT that is approximately optimal w.r.t. distributions $\{\widehat{D}_{ij}\}_{i\in[n],j\in[m]}$. We show that the computed simple mechanism is approximately optimal for the true distributions as well. The proof of Theorem 4 is postponed to Appendix F.

Theorem 4. *Suppose all bidders' valuations are constrained additive. If for each $i \in [n]$ and $j \in [m]$, D_{ij} is supported on numbers in $[0,1]$ with bit complexity no more than b, then for any $\varepsilon > 0$ and $\delta > 0$, with probability $1-\delta$, we can compute in time poly$(n, m, 1/\varepsilon, \log(1/\delta), b)$ a rationed posted price mechanism or a two-part tariff mechanism, whose revenue is at least $c \cdot \text{OPT} - O(nm^2\varepsilon)$ for some absolute constant c. The algorithm takes $O\left(\frac{\log(nm/\delta)}{\varepsilon^2}\right)$ samples from each D_{ij} and assumes query access to a demand oracle for each bidder.*

[16]This corresponds to the fact that in Lemma 1 (and Figure 1), there is a single c_i that represents δ_{ij}.

A Additional Preliminaries

Definition 7. *[RW15] Let $\mathcal{D}_i$ be the type distribution of bidder i and denote by $\mathcal{V}_i$ her distribution over valuations $v_i(t_i,\cdot)$ where $t_i \sim \mathcal{D}_i$. We say that $\mathcal{V}_i$ is subadditiver over independent items if*

- *$v_i(\cdot,\cdot)$ has no externalities, that is for any $S \subseteq [m]$, $t_i, t_i' \in \mathcal{T}_i$ such that $t_{ij} = t_{ij}'$ for $j \in S$, then $v_i(t_i,S) = v_i(t_i',S)$.*
- *$v_i(\cdot,\cdot)$ is monotone, that is for each $t_i \in \mathcal{T}_i$ and $S \subseteq T \subseteq [m]$, $v_i(t_i,S) \leq v_i(t_i,T)$*
- *$v_i(\cdot,\cdot)$ is subadditive function, that is for all $t_i \in \mathcal{T}_i$ and $S_1, S_2 \subseteq [m]$, $v_i(t_i, S_1 \cup S_2) \leq v_i(t_i,S_1) + v_i(t_i,S_2)$*

Mechanisms: A mechanism M in multi-item auctions can be described as a tuple (x,p). For every type profile t, buyer i and bundle $S \subseteq [m]$, $x_{iS}(t)$ is the probability of buyer i receiving the exact bundle S at profile t, $p_i(t)$ is the payment for buyer i at the same type profile. To ease notations, for every buyer i and types t_i, we use $p_i(t_i) = \mathbb{E}_{t_{-i}}[p_i(t_i,t_{-i})]$ as the interim price paid by buyer i and $\sigma_{iS}(t_i) = \mathbb{E}_{t_{-i}}[x_{iS}(t_i,t_{-i})]$ as the interim probability of receiving the exact bundle S.

IC and IR constraints: A mechanism $M = (x,p)$ is BIC if:

$$\sum_{S\subseteq[m]} \sigma_{iS}(t_i)\cdot v_i(t_i,S) - p_i(t_i) \geq \sum_{S\subseteq[m]} \sigma_{iS}(t_i')\cdot v_i(t_i,S) - p_i(t_i'), \forall i, t_i, t_i' \in \mathcal{T}_i.$$

The mechanism is DSIC if:

$$\sum_{S\subseteq[m]} x_{iS}(t_i,t_{-i})\cdot v_i(t_i,S) - p_i(t_i,t_{-i}) \geq \sum_{S\subseteq[m]} x_{iS}(t_i',t_{-i})\cdot v_i(t_i,S) - p_i(t_i',t_{-i}), \forall i, t_i, t_i' \in \mathcal{T}_i, t_{-i} \in \mathcal{T}_{-i}.$$

The mechanism is (interim) IR if:

$$\sum_{S\subseteq[m]} \sigma_{iS}(t_i)\cdot v_i(t_i,S) - p_i(t_i) \geq 0, \forall i, t_i \in \mathcal{T}_i.$$

The mechanism is ex-post IR if:

$$\sum_{S\subseteq[m]} x_{iS}(t_i,t_{-i})\cdot v_i(t_i,S) - p_i(t_i,t_{-i}) \geq 0, \forall i, t_i \in \mathcal{T}_i, t_{-i} \in \mathcal{T}_{-i}.$$

Definition 8 (Separation Oracle for Convex Polytope $\mathcal{P}$)**.** *A Separation Oracle SO for a convex polytope $\mathcal{P} \subseteq \mathbb{R}^d$, takes as input a point $\boldsymbol{x} \in \mathbb{R}^d$ and if $\boldsymbol{x} \in \mathcal{P}$, then the oracle says that the point is in the polytope. If $\boldsymbol{x} \notin \mathcal{P}$, then the oracle output a separating hyperplane, that is it outputs a vector $\mathbf{y} \in \mathbb{R}^d$ and $c \in \mathbb{R}$ such that $\mathbf{y}^T x \leq c$, but for $\mathbf{z} \in \mathcal{P}$, $\mathbf{y}^T\mathbf{z} > c$.*

Definition 9 (Polytopes and Facet-Complexiy)**.** *We say P has* facet-complexity *at most b if it can be written as $P := \{\vec{x} \mid \vec{x}\cdot\vec{w}^{(i)} \leq c_i,\ \forall i \in \mathcal{I}\}$, where each $\vec{w}^{(i)}$ and c_i has bit complexity at most b for all $i \in \mathcal{I}$. We use the term* convex polytope *to refer to a set of points that is closed, convex, bounded,[17] and has finite facet-complexity.*

Definition 10 (Vertex-Complexity)**.** *We use the term* corner *to refer to non-degenerate extreme points of a convex polytope. In other words, $\vec{y}$ is a corner of the d-dimensional convex polytope P if $\vec{y} \in P$ and there exist d linearly independent directions $\vec{w}^{(1)},\ldots,\vec{w}^{(d)}$ such that $\vec{x}\cdot\vec{w}^{(i)} \leq \vec{y}\cdot\vec{w}^{(i)}$ for all $\vec{x} \in P, 1 \leq i \leq d$. We use* **CORNER**$(P)$ *to denote the set of corners of a convex polytope P. We say P has vertex-complexity at most b if all vectors in* CORNER(P) *have bit complexity no more than b.*

[17] $P \subseteq \mathbb{R}^d$ is bounded if it is contained in $[-x,x]^d$ for some $x \in \mathbb{R}$.

The following fact states that the vertex-complexity and facet-complexity of a polytope in $\mathbb{R}^d$ are off by at most a d^2 multiplicative factor.

Fact 1 (Lemma 6.2.4 of [GLS12])**.** *Let P be a convex polytope in $\mathbb{R}^d$. If P has facet-complexity at most b, its vertex-complexity is at most $O(b \cdot d^2)$. Similarly, if P has vertex-complexity at most ℓ, its facet-complexity is at most $O(\ell \cdot d^2)$.*

Theorem 5. *[Ellipsoid Algorithm for Linear Programming [GLS12][18]] Let P be a convex polytope in $\mathbb{R}^d$ specified via a separation oracle SO, and $\vec{c}$ is any fixed vector in $\mathbb{R}^d$. Assume that P's facet-complexity and the bit complexity of $\vec{c}$ are no more than b. Then we can run the ellipsoid algorithm to optimize $\vec{c} \cdot \vec{x}$ over P, maintaining the following properties:*

1. *The algorithm will only query SO on rational points with bit complexity poly(d, b).*
2. *The algorithm will solve the linear program in time poly$(d, b, RT_{SO}(\text{poly}(d, b)))$, where $RT_{SO}(x)$ is the running time of the SO on any input of bit complexity x.*
3. *The output optimal solution is a corner of P.*

B Some Examples

B.1 Non-Concavity of CORE

In this section, we show that the $\text{CORE}(\sigma, \theta(\sigma))$ function is non-concave in the interim allocation rule σ. We first provide the formal definition of $\theta(\sigma)$ for a single-bidder two-item instance, and we use $\text{CORE}^{CZ}(\sigma)$ to denote $\text{CORE}(\sigma, \theta(\sigma))$.

Definition 11 (Core for a single additive bidder over two items with continuous distributions - [CZ17])**.** *Consider a single bidder interested in two items, whose value is sampled from continuous distribution D with support $T = \text{SUPP}(D)$ and density function $f(t)$ for $t \in T$. Consider a feasible interim allocation $\sigma = \{\sigma_1(t), \sigma_2(t)\}_{t \in \text{SUPP}(D)}$, that is $\sigma_1(t)$ ($\sigma_2(t)$ resp.) is the probability that the allocation rule awards item 1 (item 2 resp.) to a bidder with type t. Define*

$$\beta_1(\sigma) = \arg\min_{\beta \geq 0} \left[\Pr_{t_1 \sim D_1}[t_1 \geq \beta] = \mathbb{E}_{t \sim D}[\sigma_1(t)] \right] \qquad \beta_2(\sigma) = \arg\min_{\beta \geq 0} \left[\Pr_{t_2 \sim D_2}[t_2 \geq \beta] = \mathbb{E}_{t \sim D}[\sigma_2(t)] \right]$$

and

$$c(\sigma) = \arg\min_{a \geq 0} \left\{ \Pr_{t \sim D}[t_1 \leq \beta_1(\sigma) + a] + \Pr_{t \sim D}[t_2 \leq \beta_2(\sigma) + a] \geq \frac{1}{2} \right\}$$

The term CORE^{CZ} for interim allocation σ is defined as follows:

$$\text{CORE}^{CZ}(\sigma) = \mathbb{E}_{t \sim D}\left[\sigma_1(t) t_1 \cdot \mathbb{1}[t_1 \leq \beta_1(\sigma) + c(\sigma)]\right] + \mathbb{E}_{t \sim D}\left[\sigma_2(t) t_2 \cdot \mathbb{1}[t_2 \leq \beta_2(\sigma) + c(\sigma)]\right]$$

In Example 1 we show that $\text{CORE}^{CZ}(\sigma)$ is a non-concave function even in the setting with a single bidder and two items. The reason for the CORE^{CZ} being non-concave lies in the fact that the interval which we truncate depends on the interim allocation σ. Computing the concave hull of $\text{CORE}^{CZ}(\sigma)$ in the worst case requires exponential time in the dimension of the space, which is m is our case.

Example 1. *Consider a single additive bidder interested in two items whose values are both drawn from the uniform distribution $U[0, 1]$. Consider two interim allocation rules σ and σ':*

[18] Properties 1 and 2 follow from Theorem 6.4.9 of [GLS12] and Property 3 follows from Remark 6.5.2 of [GLS12].

- *σ: Award the first item to the buyer if her value for it lies in the interval $[0, 1/2]$ and never award the second item to the buyer.*

- *σ': Always award the first item to the buyer and never award the second item to the buyer.*

According to Definition 12, for allocation rule σ, the dual parameters are $\beta_1(\sigma) = 1/2$, $\beta_2(\sigma) = 1$ and $c(\sigma) = 0$, which implies $\text{CORE}^{CZ}(\sigma) = 1/8$. Similarly for allocation rule σ' we have $\beta_1(\sigma') = 0$, $\beta_2(\sigma') = 1$ and $c(\sigma') = 0$, which implies that $\text{CORE}^{CZ}(\sigma') = 0$.

Consider the interim allocation σ'' that uses allocation rule σ with probability 50% and σ' with 50%. Note that σ'' is in the convex combination of σ and σ' and more specifically $\sigma'' = \frac{\sigma+\sigma'}{2}$. For interim allocation σ'' we have that $\beta_1(\sigma'') = 1/4$, $\beta_2(\sigma'') = 1$ and $c(\sigma'') = 0$, which implies that $\text{CORE}^{CZ}(\sigma'') = \frac{1}{32}$. We notice that the second item contributes nothing to the CORE^{CZ}, but it ensures that $c = 0$ regardless of the allocation of the first item. Thus $\text{CORE}^{CZ}\left(\sigma''\right) < \frac{1}{2}(\text{CORE}^{CZ}(\sigma) + \text{CORE}^{CZ}(\sigma'))$, which implies that $\text{CORE}^{CZ}(\cdot)$ is not a concave function.

B.2 Why can't we use the Ex-Ante Relaxation?

An influential framework known as the ex-ante relaxation has been widely used in Mechanism Design, but is insufficient for our problem. Informally speaking, ex-ante relaxation reduces a multi-bidder objective to the sum of single-bidder objectives subject to some global supply constraints over ex-ante allocation probabilities. To solve the ex-ante relaxation program efficiently, the single-bidder objective has to be concave and efficiently computable given the ex-ante probabilities [Ala11].

In revenue maximization, the single-bidder objective – the optimal revenue subject to ex-ante probabilities – is indeed a concave function. However, we do not have a polynomial time algorithm to even compute the single-bidder objective given a set of fixed ex-ante probabilities.[19] To fix this issue, one can try to find a concave function that is always a good approximation to the single-bidder objective for any ex-ante probabilities. To the best of our knowledge, such a concave function only exists for unit-demand bidders via the copies setting technique [CHMS10b]. Alternatively, one can replace the global objective – optimal revenue by the upper bound of revenue proposed in [CZ17]. Yet the corresponding single-bidder objective for one term CORE in the upper bound is highly non-concave, which makes the ex-ante relaxation not applicable.

Although the term CORE was originally defined for interim allocation rules (as in Definition 11), it can also be defined for ex-ante probabilities. We only define it for the single-bidder two-item case. Let $q = \{q_1, q_2\} \in [0, 1]^2$, and $\text{MAX-CORE} = \max_{\sigma\in\Sigma(q)} \text{CORE}^{CZ}(\sigma)$, where $\Sigma(q)$ is the set of feasible interim allocations that awards the first item with probability at most q_1 and the second item with probability at most q_2. Example 2 also shows that $\text{MAX-CORE}(\cdot)$ is a non-concave function by observing that $\sigma \in \Sigma(1/2, 0)$,$\sigma' \in \Sigma(1, 0)$ and $\sigma'' \in \Sigma(3/4, 0)$.

Definition 12 (Core for a single additive bidder over two items - [CZ17])**.** *Consider a single bidder interested in two items, whose value is sampled from $D_1 \times D_2$. Consider a supply constraints $q_1, q_2 \in [0, 1]$. Note that q_1 (or q_2) is the probability that a mechanism awards the first item (or the second item) to the bidder. Define*

$$\beta_1 = \arg\min_{\beta\geq 0}\left[\Pr_{t_1\sim D_1}[t_1 \geq \beta] = q_1\right] \qquad \beta_2 = \arg\min_{\beta\geq 0}\left[\Pr_{t_2\sim D_2}[t_2 \geq \beta] = q_2\right]$$

and

$$c = \arg\min_{a\geq 0}\left\{\Pr_{t_1\sim D_1}[t_1 \leq \beta_1 + a] + \Pr_{t_2\sim D_2}[t_2 \leq \beta_2 + a] \geq \frac{1}{2}\right\}$$

[19] The closest thing we know is a QPTAS for a unit-demand bidder. See Section 1.2.

The term MAX-CORE *is defined as follows:*

$$\text{MAX-CORE}(q) = \max_{\substack{x_1:\mathcal{T}_1\to[0,1]\\ \sum_{t_1\in\mathcal{T}_1} f_1(t_1)x_1(t_1)=q_1}} \sum_{\substack{t_1\in\mathcal{T}_1\\ t_1\le\beta_1+c}} f_1(t_1)\cdot t_1\cdot x_1(t_1) + \max_{\substack{x_2:\mathcal{T}_2\to[0,1]\\ \sum_{t_2\in\mathcal{T}_2} f_2(t_2)x_2(t_2)=q_2}} \sum_{\substack{t_2\in\mathcal{T}_2\\ t_2\le\beta_2+c}} f_2(t_2)\cdot t_2\cdot x_2(t_2)$$

In Example 2 we show that MAX-CORE(q) is a non-concave function even in the setting with a single bidder and two items. The reason for the MAX-CORE being non-concave lies in the fact that the interval which we truncate depends on the supply constraints q. Computing the concave hull of MAX-CORE(q) in the worst case requires exponential time in the dimension of the space, which is m is our case. These facts make the ex-ante relaxation approach not applicable to solve our problem.

Example 2. *Consider a single additive bidder interested in two items whose values are both drawn from the uniform distribution $U[0,1]$. Consider the values $q = (1/2, 0)$ and $q' = (1, 0)$. According to Definition 12, for q we have that $\beta_1^{(q)} = 1/2, \beta_2^{(q)} = 1$ and $c^{(q)} = 0$ and for q' we have $\beta_1^{(q')} = 0, \beta_2^{(q')} = 1$ and $c^{(q')} = 0$. We notice that the second item contributes nothing to the* MAX-CORE*, but it ensures that $c = 0$ regardless of the supply demand for the first item. Observe that* MAX-CORE$(q) = 1/8$ *and* MAX-CORE$(q') = 0$. *Let $q'' = (q + q')/2 = (3/4, 0)$. For q'', observe that $\beta_1^{(q'')} = 1/4, \beta_2^{(q'')} = 1$ and $c^{(q'')} = 0$. We have* MAX-CORE$(q'') = 1/32$. *Thus* MAX-CORE $\left(q''\right) < \frac{1}{2}($MAX-CORE$(q) +$ MAX-CORE$(q'))$, *which implies that* MAX-CORE$(\cdot)$ *is not a concave function.*

C Missing Details from Section 3

In this section, we provide a proof of Theorem 3. Indeed, we prove a generalization that works for XOS buyers (Theorem 6), with the generalized of the single-bidder marginal reduced form polytope Definition 14 and a generalized LP (Figure 4).

Theorem 6. *Let $(w, \lambda, \hat{\lambda}, \boldsymbol{d})$ (or $(\pi, w, \lambda, \hat{\lambda}, \boldsymbol{d})$) be any feasible solution of the LP in Figure 3 (or Figure 4). Let $\mathcal{M}_{PP}$ be the rationed posted price mechanism computed in Lemma 2. Let $\mathcal{M}_{TPT}$ be the two-part tariff mechanism shown in Mechanism 1 with prices $\{Q_j\}_{j\in[m]}$ (Definition 15). Then the objective function of the solution $2 \cdot \sum_{j\in[m]} Q_j$ is bounded by $c_1 \cdot \text{REV}(\mathcal{M}_{PP}) + c_2 \cdot \text{REV}(\mathcal{M}_{TPT})$, for some constant $c_1, c_2 > 0$. Moreover, both $\mathcal{M}_{PP}$ and $\mathcal{M}_{TPT}$ can be computed in time poly$(n, m, \sum_{i,j} |\mathcal{T}_{ij}|)$, with access to the demand oracle for the buyers' valuations.*

C.1 Result by Cai and Zhao [CZ17] for XOS Valuations

The result by Cai and Zhao [CZ17] applies also to XOS valuations. Here we state their result for this general case. Note that this is a generalized definition and lemma for Definition 4 and Lemma 1.

Definition 13. *For any $i \in [n], j \in [m]$, let $\mathcal{V}_{ij}^0 = \{V_{ij}(t_{ij}) : t_{ij} \in \mathcal{T}_{ij}\}$. For any feasible interim allocation σ, and non-negative numbers $\tilde{\boldsymbol{\beta}} = \{\tilde{\beta}_{ij} \in \mathcal{V}_{ij}^0\}_{i\in[n],j\in[m]}$, $\mathbf{c} = \{c_i\}_{i\in[n]}$ and $\mathbf{r} = \{r_{ij}\}_{i\in[n],j\in[m]} \in [0,1]^{nm}$ (which we refer to as the dual parameters), define* CORE$(\sigma, \tilde{\boldsymbol{\beta}}, \mathbf{c}, \mathbf{r})$ *as the welfare under allocation σ truncated at $\tilde{\beta}_{ij} + c_i$ for every i, j. Formally,*

$$\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}, \mathbf{c}, \mathbf{r}) = \sum_i \sum_{t_i} f_i(t_i) \cdot \sum_{S\subseteq[m]} \sigma_{iS}(t_i) \sum_{j\in S} t_{ij} \cdot \left(\mathbb{1}[t_{ij} < \tilde{\beta}_{ij} + c_i] + r_{ij} \cdot \mathbb{1}[t_{ij} = \tilde{\beta}_{ij} + c_i]\right)$$

if the buyers have constrained-additive valuations, and

$$\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}, \mathbf{c}, \mathbf{r}) = \sum_i \sum_{t_i} f_i(t_i) \cdot \sum_{S\subseteq[m]} \sigma_{iS}(t_i) \sum_{j\in S} \gamma_{ij}^S(t_i) \cdot \left(\mathbb{1}[V_{ij}(t_{ij}) < \tilde{\beta}_{ij} + c_i] + r_{ij}\mathbb{1}[V_{ij}(t_{ij}) = \tilde{\beta}_{ij} + c_i]\right)$$

if the buyers have XOS valuations. Here $\gamma_{ij}^S(t_i) = \alpha_{ij}^{k^*(t_i,S)}(t_{ij})$, where $k^*(t_i, S) = \arg\max_{k\in[K]} \left(\sum_{j\in S} \alpha_{ij}^k(t_{ij})\right)$.

Lemma 4. *[CZ17] Given any BIC and IR mechanism $\mathcal{M}$, there exists **(i)** a feasible interim allocation σ,*[20] *where $\sigma_{iS}(t_i)$ is the interim probability for buyer i to receive exactly bundle S when her type is t_i, **(ii)** non-negative numbers $\tilde{\boldsymbol{\beta}}^{(\sigma)} = \{\tilde{\beta}_{ij}^{(\sigma)} \in \mathcal{V}_{ij}^0\}_{i\in[n],j\in[m]}$, $\mathbf{c}^{(\sigma)} = \{c_i^{(\sigma)}\}_{i\in[n]}$ and $\mathbf{r}^{(\sigma)} \in [0,1]^{nm}$ that depend on σ, and **(iii)** a two-part tariff mechanism $\mathcal{M}_1^{(\sigma)}$ such that*

1. $\sum_{i\in[n]} \left(\Pr_{t_{ij}}[V_{ij}(t_{ij}) > \tilde{\beta}_{ij}^{(\sigma)}] + r_{ij}^{(\sigma)} \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) = \tilde{\beta}_{ij}^{(\sigma)}]\right) \le \frac{1}{2}$.
2. $\frac{1}{2} \cdot \sum_{t_i\in\mathcal{T}_i} f_i(t_i) \cdot \sum_{S:j\in S} \sigma_{iS}(t_i) \le \Pr_{t_{ij}}[V_{ij}(t_{ij}) > \tilde{\beta}_{ij}^{(\sigma)}] + r_{ij}^{(\sigma)} \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) = \tilde{\beta}_{ij}^{(\sigma)}], \forall i, j.$
3. $\textsc{Rev}(\mathcal{M}) \le 28 \cdot \textsc{PRev} + 4 \cdot \textsc{Core}(\sigma, \tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)}, \mathbf{r}^{(\sigma)})$.
4. $\sum_{i\in[n]} c_i^{(\sigma)} \le 8 \cdot \textsc{PRev}$.
5. $\textsc{Core}(\sigma, \tilde{\boldsymbol{\beta}}^{(\sigma)}, \mathbf{c}^{(\sigma)}, \mathbf{r}^{(\sigma)}) \le 64 \cdot \textsc{PRev} + 8 \cdot \textsc{Rev}(\mathcal{M}_1^{(\sigma)})$.

C.2 Single-Bidder Marginal Reduced Form Polytope for XOS Valuations

In Definition 14 we define the single-bidder marginal reduced form polytope W_i for XOS buyers, which differs from the single-bidder marginal reduced form polytope for constrained-additive valuation is several ways. In Definition 14, we define a distribution σ_S^k over all possible subset of items $S \subseteq [m]$ and over the finite number $k \in [K]$ over additive functions that can be chosen when we evaluate the value that the buyer has for a set of items. In Definition 5, the distribution σ_S was only over sets in the set of feasible allocations.

Similar to Definition 5, $\pi_{ij}(t_{ij})$ is equal to $f_{ij}(t_{ij})$ times the probability that the i-th buyer receives the j-th item. In contrast to Definition 5, in Definition 14, the value of $w_{ij}(t_{ij})$ is $\frac{f_{ij}(t_{ij})}{V_{ij}(t_{ij})}$ times the expected value that the buyer has for the item when we are allowed to choose which additive functions in $k \in [K]$ we count the value of the buyer, or we are even allowed to allocate an item to the buyer but count zero value for it (that is equivalent to just throwing away the item).

Definition 14 (XOS valuations: single-bidder marginal reduced form polytope)**.** *For every $i \in [n]$, the single-bidder marginal reduced form polytope $W_i \subseteq [0,1]^{2\cdot\sum_j |\mathcal{T}_{ij}|}$ is defined as follows. Let $\pi_i = (\pi_{ij}(t_{ij}))_{j,t_{ij}\in\mathcal{T}_{ij}}$ and $w_i = (w_{ij}(t_{ij}))_{j,t_{ij}\in\mathcal{T}_{ij}}$. Then $(\pi_i, w_i) \in W_i$ if and only if there exist a number $\sigma_S^{(k)}(t_i) \in [0,1]$ for every $t_i \in \mathcal{T}_i, S \subseteq [m], k \in [K]$, such that*

1. $\sum_{S,k} \sigma_S^{(k)}(t_i) \le 1$, $\forall t_i \in \mathcal{T}_i$.
2. $\pi_{ij}(t_{ij}) = f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S:j\in S} \sum_{k\in[K]} \sigma_S^{(k)}(t_{ij}, t_{i,-j})$, *for all* $i, j, t_{ij} \in \mathcal{T}_{ij}$.
3. $w_{ij}(t_{ij}) \le f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S:j\in S} \sum_{k\in[K]} \sigma_S^{(k)}(t_{ij}, t_{i,-j}) \cdot \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})}$, *for all* $i, j, t_{ij} \in \mathcal{T}_{ij}$.

[20] Note that when buyers have constrained-additive valuations, it suffice to take σ to be the interim allocation rule of $\mathcal{M}$. For XOS valuations, σ will be the interim allocation of a modified version of $\mathcal{M}$. See Section 5 in [CZ16] for details.

C.3 The Linear Program for XOS valuations

$$\textbf{max} \sum_{i\in[n]}\sum_{j\in[m]}\sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij})\cdot V_{ij}(t_{ij})\cdot \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta}} \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\cdot \mathbb{1}[V_{ij}(t_{ij})\le \beta_{ij}+\delta_{ij}]$$

s.t.

Allocation Feasibility Constraints:

(1) $(\pi_i, w_i)\in W_i \quad \forall i$

(2) $\sum_i \sum_{t_{ij}\in\mathcal{T}_{ij}} \pi_{ij}(t_{ij})\le 1 \quad \forall j$

Natural Feasibility Constraints:

(3) $f_{ij}(t_{ij})\cdot \sum_{\beta_{ij}\in\mathcal{V}_{ij}}\sum_{\delta_{ij}\in\Delta}\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij}) = w_{ij}(t_{ij}) \quad \forall i,j,t_{ij}\in\mathcal{T}_{ij}$

(4) $\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\le \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij}) \quad \forall i,j,t_{ij},\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}$

(5) $\sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta}} \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij}) = 1 \quad \forall i,j$

Problem Specific Constraints:

(6) $\sum_{i\in[n]}\sum_{\beta_{ij}\in\mathcal{V}_{ij}}\sum_{\delta_{ij}\in\Delta}\hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\cdot \Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij})\ge\beta_{ij}]\le\frac{1}{2} \quad \forall j$

(7) $\frac{1}{2}\sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij})\left(\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})+\lambda_{ij}(t_{ij},\beta_{ij}^+,\delta_{ij})\right)\le$

$\hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\cdot \Pr_{t_{ij}}[V_{ij}(t_{ij})\ge\beta_{ij}] + \hat{\lambda}_{ij}(\beta_{ij}^+,\delta_{ij})\cdot\Pr_{t_{ij}}[V_{ij}(t_{ij})\ge\beta_{ij}^+] \quad \forall i,j,\beta_{ij}\in\mathcal{V}_{ij}^0,\delta_{ij}\in\Delta$

(8) $\sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta}} \delta_{ij}\cdot\hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\le d_i \quad \forall i,j$

(9) $\sum_{i\in[n]} d_i \le 111\cdot \widetilde{\text{PREV}}$

$\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\ge 0, \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij})\ge 0, \pi_{ij}(t_{ij})\ge 0, w_{ij}(t_{ij})\ge 0, d_i\ge 0 \quad \forall i,j,t_{ij},\beta_{ij}\in\mathcal{V}_{ij},\delta_{ij}$

Figure 4: LP for XOS Valuations

The LP for XOS valuations can be found in Figure 4. Here $V_{ij}^0=\{V_{ij}(t_{ij}): t_{ij}\in\mathcal{T}_{ij}\}$, $\mathcal{V}_{ij}^+=\{V_{ij}(t_{ij})+\varepsilon_r : t_{ij}\in\mathcal{T}_{ij}\}$ and $\mathcal{V}_{ij}=\mathcal{V}_{ij}^0\cup\mathcal{V}_{ij}^+$. We notice that this is consistent with our LP for constrained-additive buyers (Figure 3), as $V_{ij}(t_{ij})=t_{ij}$ for constrained-additive buyers.

Denote OPT_{LP} the optimum objective of the LP in Figure 4. Similar to the constrained-additive case, we have the following lemma.

Lemma 5. *When buyers have XOS valuations, for any BIC and IR mechanism $\mathcal{M}$,* $\text{REV}(\mathcal{M})\le 28\cdot\text{PREV}+4\cdot\text{OPT}_{\text{LP}}$.

C.4 Proof of Lemma 3 and Lemma 5

Proof. The proof is stated for XOS buyers, whose LP contains a new set of variables π compared to the LP for constrained-additive buyers. When the buyers have constrained-additive valuations, we can simply treat π to be the same as w. Also, note that $V_{ij}(t_{ij})=t_{ij}$ for constrained-additive valuations.

Let tuple $(\hat{\sigma}, \tilde{\boldsymbol{\beta}}, \mathbf{c}, \mathbf{r})$ be the one stated in Lemma 4 for $\mathcal{M}$. Consider the following choice of variables of the LP in Figure 3 (or Figure 4). For every i, j, t_{ij}, let

$$w_{ij}(t_{ij}) = f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j} \in \mathcal{T}_{i,-j}} f_{i,-j}(t_{i,-j}) \sum_{S: j \in S} \hat{\sigma}_{iS}(t_{ij}, t_{i,-j})$$

if the buyers' have constrained-additive valuations. Let

$$\pi_{ij}(t_{ij}) = f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j} \in \mathcal{T}_{i,-j}} f_{i,-j}(t_{i,-j}) \sum_{S: j \in S} \hat{\sigma}_{iS}(t_{ij}, t_{i,-j})$$

$$w_{ij}(t_{ij}) = f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j} \in \mathcal{T}_{i,-j}} f_{i,-j}(t_{i,-j}) \sum_{S: j \in S} \hat{\sigma}_{iS}(t_{ij}, t_{i,-j}) \cdot \frac{\gamma_{ij}^S(t_i)}{V_{ij}(t_{ij})}.$$

if the buyers' have XOS valuation. For each c_i, we notice that by Lemma 4 and Lemma 2, $0 \leq c_i \leq 8 \cdot \text{PREV} \leq 55 \cdot \widetilde{\text{PREV}}$ when $n \cdot m \geq 110$. We round it up to the closest number in Δ, and we denote it using $\hat{c}_i$. Clearly,

$$\text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}, \mathbf{c}, \mathbf{r}) \leq \text{CORE}(\sigma, \tilde{\boldsymbol{\beta}}, \hat{\mathbf{c}}, \mathbf{r}).$$

$\lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij})$ and $\hat{\lambda}_{ij}(\beta_{ij}, \delta_{ij})$ can be set to non-zero only if $\beta_{ij} \in \{\tilde{\beta}_{ij}, \tilde{\beta}_{ij}^+\}$ and $\delta_{ij} = \hat{c}_i$. More specifically, we choose the variables as follows.

- $\lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}, c_i) := r_{ij} \cdot w_{ij}(t_{ij}) / f_{ij}(t_{ij})$,
- $\lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}^+, c_i) := (1 - r_{ij}) \cdot w_{ij}(t_{ij}) / f_{ij}(t_{ij})$,
- $\hat{\lambda}_{ij}(\tilde{\beta}_{ij}, c_i) = r_{ij}$,
- $\hat{\lambda}_{ij}(\tilde{\beta}_{ij}^+, c_i) = 1 - r_{ij}$;
- $d_i = \hat{c}_i$.

We show that this is indeed a feasible solution of the LP in Figure 3 by verifying each constraint. We first prove that $(\pi_i, w_i) \in W_i$ for every i. This is clear for constrained-additive valuations. For XOS valuations, consider the mapping $\sigma_{iS}^{(k)}(t_i) = \hat{\sigma}_{iS}(t_i) \cdot \mathbb{1}[k = \arg\max_{k' \in [K]} \sum_{j \in S} \alpha_{ij}^{(k')}(t_{ij})]$ for every t_i (we break ties arbitrarily). Thus by the definition of γ_{ij}^S, we have $\hat{\sigma}_{iS}(t_i) \cdot \gamma_{ij}^S(t_i) = \sum_k \sigma_{iS}^{(k)}(t_i) \cdot \alpha_{ij}^{(k)}(t_i)$. Then clearly (π_i, w_i) satisfies all of the conditions in Definition 14.

For Constraint (3), LHS equals to $f_{ij}(t_{ij}) \cdot \left(\lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}, c_i) + \lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}^+, c_i)\right) = w_{ij}(t_{ij})$. Constraint (2) follows from the fact that $\hat{\sigma}$ is a feasible interim allocation rule, and each item j can be allocated to at most one buyer for every type profile. Thus

$$\sum_i \sum_{t_{ij}} \pi_{ij}(t_{ij}) = \sum_i \sum_{t_i} f_i(t_i) \sum_{S: j \in S} \hat{\sigma}_{iS}(t_i) \leq 1$$

By property 1 of Lemma 4 and the choice of ε_r, we have that for every j,

$$\begin{aligned}
&\sum_{i \in [n]} \sum_{\beta_{ij} \in \mathcal{V}_{ij}} \hat{\lambda}_{ij}(\beta_{ij}) \cdot \Pr_{t_{ij} \sim D_{ij}}[V_{ij}(t_{ij}) \geq \beta_{ij}] \\
=& \sum_{i \in [n]} \left(r_{ij} \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) \geq \tilde{\beta}_{ij}] + (1 - r_{ij}) \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) \geq \tilde{\beta}_{ij}^+] \right) \\
=& \sum_{i \in [n]} \left(\Pr_{t_{ij}}[V_{ij}(t_{ij}) > \tilde{\beta}_{ij}] + r_{ij} \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) = \tilde{\beta}_{ij}] \right) \leq \frac{1}{2}
\end{aligned}$$

Thus, Constraint (6) is satisfied. For Constraint (7), we only need to verify the constraint for $\beta_{ij} = \tilde{\beta}_{ij} \in \mathcal{V}_{ij}^0$. LHS equals to $\frac{1}{2}\sum_{t_{ij}} w_{ij}(t_{ij})$. We notice that $\gamma_{ij}^S(t_i) \leq V_{ij}(t_{ij})$ for all i, j, S (for XOS valuations). Thus

$$\sum_{t_{ij}\in\mathcal{T}_{ij}} w_{ij}(t_{ij}) \leq \sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij}) \cdot \left(\sum_{t_{i,-j}\in\mathcal{T}_{i,-j}} f_{i,-j}(t_{i,-j}) \sum_{S:j\in S} \hat{\sigma}_{iS}(t_{ij}, t_{i,-j})\right) = \sum_{t_i\in\mathcal{T}_i} f_i(t_i) \cdot \sum_{S:j\in S} \hat{\sigma}_{iS}(t_i)$$

Since $r_{ij} \in [0,1]$ for all i, j, by property 2 of Lemma 4 and the choice of ε_r, we have

$$\begin{aligned}\text{LHS of Constraint (7)} \leq & \frac{1}{2}\sum_{t_i\in\mathcal{T}_i} f_i(t_i) \cdot \sum_{S:j\in S} \hat{\sigma}_{iS}(t_i) = \Pr_{t_{ij}}[V_{ij}(t_{ij}) > \tilde{\beta}_{ij}] + r_{ij} \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) = \tilde{\beta}_{ij}] \\ = & \hat{\lambda}_{ij}(\beta_{ij}^+) \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) \geq \beta_{ij}^+] + \hat{\lambda}_{ij}(\beta_{ij}) \cdot \Pr_{t_{ij}}[V_{ij}(t_{ij}) \geq \beta_{ij}]\end{aligned}$$

For Constraint (4), since $\gamma_{ij}^S(t_i) \leq V_{ij}(t_{ij})$ for all i, j, S and $\sum_{S\in[m]} \hat{\sigma}_{iS}(t_i) \leq 1$, we have $w_{ij}(t_{ij}) \leq f_{ij}(t_{ij})$. Thus

$$\lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}, c_i) = r_{ij} \cdot \frac{w_{ij}(t_{ij})}{f_{ij}(t_{ij})} \leq r_{ij} = \hat{\lambda}_{ij}(\tilde{\beta}_{ij}, c_i).$$

and

$$\lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}^+, c_i) = (1 - r_{ij}) \cdot \frac{w_{ij}(t_{ij})}{f_{ij}(t_{ij})} \leq 1 - r_{ij} = \hat{\lambda}_{ij}(\tilde{\beta}_{ij}^+, c_i).$$

Constraint (8) and (5) are straightforward since $\hat{\lambda}_{ij}(\beta_{ij}, \delta_i) = r_{ij} \cdot \mathbb{1}[\beta_{ij} = \tilde{\beta}_{ij} \wedge \delta_{ij} = \hat{c}_i]$, $\hat{\lambda}_{ij}(\beta_{ij}, \delta_i) = (1 - r_{ij}) \cdot \mathbb{1}[\beta_{ij} = \tilde{\beta}_{ij}^+ \wedge \delta_{ij} = \hat{c}_i]$, and $d_i = \hat{c}_i$.

Lastly, for Constraint (9), we notice that for every $i \in [n]$, $\hat{c}_i \leq \max\{\frac{\widetilde{\text{PREV}}}{n}, 2c_i\}$. Thus by property 4 of Lemma 4, when $n \cdot m \geq 110$,

$$\sum_{i\in[n]} c_i \leq 2 \cdot \sum_{i\in[n]} c_i + \widetilde{\text{PREV}} \leq 16 \cdot \text{PREV} + \widetilde{\text{PREV}} \leq 111 \cdot \widetilde{\text{PREV}}$$

Thus the solution is feasible. We are left to show that the objective of the above solution is at least $\textsc{Core}(\hat{\sigma}, \tilde{\boldsymbol{\beta}}, \hat{\mathbf{c}}, \mathbf{r})$. In fact, by the definition of $w_{ij}(t_{ij})$,

$$\begin{aligned}\textsc{Core}(\hat{\sigma}, \tilde{\boldsymbol{\beta}}, \hat{\mathbf{c}}, \mathbf{r}) = & \sum_{i\in[n]}\sum_{j\in[m]}\sum_{t_{ij}\in\mathcal{T}_{ij}} w_{ij}(t_{ij}) \cdot V_{ij}(t_{ij}) \cdot \Big(\mathbb{1}[V_{ij}(t_{ij}) < \tilde{\beta}_{ij} + \hat{c}_i] + r_{ij} \cdot \mathbb{1}[V_{ij}(t_{ij}) = \tilde{\beta}_{ij} + \hat{c}_i]\Big) \\ \leq & \sum_i\sum_j\sum_{t_{ij}} w_{ij}(t_{ij}) \cdot V_{ij}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \tilde{\beta}_{ij} + \hat{c}_i].\end{aligned}$$

This is exactly the objective of the LP in Figure 3 according to the choice of ε_r, since

$$\begin{aligned}\lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}, \delta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \tilde{\beta}_{ij} + \hat{c}_i] + \lambda_{ij}(t_{ij}, \tilde{\beta}_{ij}^+, \delta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \tilde{\beta}_{ij}^+ + \hat{c}_i] \\ = \frac{w_{ij}(t_{ij})}{f_{ij}(t_{ij})} \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \tilde{\beta}_{ij} + \hat{c}_i]\end{aligned}$$

The proof is complete by invoking property 3 of Lemma 4. □

C.5 Bounding the Difference between the Shifted CORE and the Original CORE

We first give the following definition as a generalization of Definition 6.

Definition 15. *Let $(\pi, w, \lambda, \hat{\lambda}, \boldsymbol{d} = (d_i)_{i\in[n]})$ be any feasible solution of the LP in Figure 4. For every $j \in [m]$, define*

$$Q_j = \frac{1}{2} \cdot \sum_{i\in[n]} \sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij}) \cdot V_{ij}(t_{ij}) \cdot \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij} \\ \delta_{ij}\in\Delta}} \lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \le \beta_{ij} + \delta_{ij}].$$

Recall that by Constraint (5), for every i, j, $\hat{\lambda}_{ij}(\cdot,\cdot)$ can be viewed as a joint distribution $\mathcal{C}_{ij}$ over $\mathcal{V}_{ij} \times \Delta$, i.e. $\Pr_{(\beta_{ij},\delta_{ij})\sim\mathcal{C}_{ij}}[\beta_{ij} = a \wedge \delta_{ij} = b] = \hat{\lambda}_{ij}(\beta_{ij} = a, \delta_{ij} = b)$. Denote $\mathcal{B}_{ij}$ the marginal distribution of β_{ij} with respect to $\mathcal{C}_{ij}$. Inspired by the "shifted core" technique by Cai and Zhao [CZ17], we need further definitions to describe the welfare contribution by each item under a smaller threshold.

Definition 16. *For every $i \in [n]$, define*[21]

$$\tau_i = \inf_{x\ge 0}\left\{\sum_{j\in[m]} \Pr_{t_{ij}\sim D_{ij},\beta_{ij}\sim\mathcal{B}_{ij}}[V_{ij}(t_{ij}) \ge \max(\beta_{ij}, Q_j + x)] \le \frac{1}{2}\right\}$$

Then for every $j \in [m]$, define

$$\hat{Q}_j = \frac{1}{2} \cdot \sum_{i\in[n]} \sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij}) \cdot V_{ij}(t_{ij}) \cdot \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij} \\ \delta_{ij}\in\Delta}} \lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \le \min\{\beta_{ij} + \delta_{ij}, Q_j + \tau_i\}]$$

We prove in Lemma 6 that the difference between $\sum_{j\in[m]} Q_j$ and $\sum_{j\in[m]} \hat{Q}_j$ can be bounded using PREV.

Lemma 6. *For every $j \in [m]$, $Q_j \ge \hat{Q}_j$. Moreover, there exists some absolute constant $c > 0$ such that*

$$\sum_{j\in[m]} Q_j \le \sum_{j\in[m]} \hat{Q}_j + c \cdot \text{PREV}$$

To prove Lemma 6, we consider the following variant of the RPP mechanism, where the posted prices are allowed to be randomized.

Rationed Randomized Posted Price Mechanism (RRPP). Before the auction starts, the seller first draws a posted price p_{ij} from some distribution $\mathcal{G}_{ij}$, for every buyer i and item j. All $\mathcal{G}_{ij}$s are independent. The buyers then arrive in some arbitrary order, and each buyer i can purchase *at most one* item among the available ones at the realized posted price p_{ij} for every item j.

Clearly any RRPP mechanism is also DSIC and IR. We notice that any RRPP mechanism can be viewed as a distribution of RPP mechanisms, as the seller can draw all the posted prices before the auction starts, and use the realized (and deterministic) set of posted prices to sell the items. Thus the highest revenue achievable among all RRPP mechanisms is the same as the optimum revenue among all RPP mechanisms, which is PREV.

Before giving the proof of Lemma 6, we first prove a useful lemma that analyzes the revenue of RRPP. It's a generalization of Lemma 17 of [CZ17], which allows randomized posted prices.

[21] If all D_{ij}s are continuous, then for every i there exists τ_i that satisfies the following property: $\sum_{j\in[m]} \Pr_{t_{ij}\sim D_{ij},\beta_{ij}\sim\mathcal{B}_{ij}}[V_{ij}(t_{ij}) \ge \max(\beta_{ij}, Q_j + x)] \le \frac{1}{2}$ and the inequality is achieved as equality for all $\tau_i > 0$. However, this property might not be satisfied for discrete distributions. This is again a tie-breaking issue addressed in Remark 1. We refer the readers to Lemma 5 of [CZ17] for a fix. For simplicity, in our proof we will assume that all τ_is satisfy the property mentioned above.

Lemma 7. *[CZ17] For every i, j, let $\mathcal{G}_{ij}$ be any distribution over $\mathbb{R}_+$. All $\mathcal{G}_{ij}$'s are independent from each other. Suppose both of the following inequalities hold, for some constant $a, b \in (0, 1)$:*

1. $\sum_{i\in[n]} \Pr_{t_{ij}\sim\mathcal{D}_{ij}, x_{ij}\sim\mathcal{G}_{ij}} \left[V_i(t_{ij}) \geq x_{ij}\right] \leq a, \forall j \in [m]$.
2. $\sum_{j\in[m]} \Pr_{t_{ij}\sim\mathcal{D}_{ij}, x_{ij}\sim\mathcal{G}_{ij}} \left[V_i(t_{ij}) \geq x_{ij}\right] \leq b, \forall i \in [n]$.

Then

$$\sum_{i\in[n]}\sum_{j\in[m]} \mathop{\mathbb{E}}_{x_{ij}\sim\mathcal{G}_{ij}}\left[x_{ij}\cdot \Pr_{t_{ij}\sim\mathcal{D}_{ij}}\left[V_i(t_{ij}) \geq x_{ij}\right]\right] \leq \frac{1}{(1-a)(1-b)}\cdot \text{PREV}.$$

Proof. Consider an RRPP that sells item j to buyer i at price $x_{ij} \sim \mathcal{G}_{ij}$. The mechanism visits the buyers in some arbitrary order. For every i, j and every realized x_{ij}, we will bound the probability of buyer i purchasing item j, over the randomness of $\{\mathcal{D}_{i'j'}\}_{i'\in[n],j'\in[m]}$ and $\{\mathcal{G}_{i'j'}\}_{(i',j')\neq(i,j)}$. Notice that when it is buyer i's turn, she purchases exactly item j and pays x_{ij} if all of the following three conditions hold: (i) j is still available, (ii) $V_i(t_{ij}) \geq x_{ij}$ and (iii) $\forall k \neq j, V_i(t_{ik}) < x_{ik}$. The second condition means buyer i can afford item j. The third condition means she cannot afford any other item $k \neq j$. Therefore, buyer i purchases exactly item j.

Now let us compute the probability that all three conditions hold, when $t_{ij} \sim \mathcal{D}_{ij}$ and $x_{ij} \sim \mathcal{G}_{ij}$ for all i, j. Since every buyer's valuation is subadditive over the items, item j is purchased by someone else only if there exists a buyer $k \neq i$ who has $V_k(t_{kj}) \geq x_{kj}$. By the union bound, the event described above happens with probability at most $\sum_{k\neq i} \Pr_{t_{kj},x_{kj}}\left[V_k(t_{kj}) \geq x_{kj}\right]$, which is less than a by Inequality 1 of the statement. Therefore, condition (i) holds with probability at least $1-a$. Clearly, condition (ii) holds with probability $\Pr_{t_{ij}}\left[V_i(t_{ij}) \geq x_{ij}\right]$. Finally, condition (iii) holds with at least probability $1-b$, because the probability that there exists any item $k \neq j$ such that $V_i(t_{ik}) \geq x_{ik}$ is no more than $\sum_{k\neq j} \Pr_{t_{ik},x_{ik}}[V_i(t_{ik}) \geq x_{ik}] \leq b$ (Inequality 2 of the statement). Since the three conditions are independent, buyer i purchases exactly item j with probability at least $(1-a)(1-b)\cdot \Pr_{t_{ij}}\left[V_i(t_{ij}) \geq x_{ij}\right]$. So the expected revenue of this mechanism is at least $(1-a)(1-b)\cdot \mathbb{E}_{x_{ij}\sim\mathcal{G}_{ij}}\left[x_{ij}\cdot \Pr_{t_{ij}\sim\mathcal{D}_{ij}}\left[V_i(t_{ij}) \geq x_{ij}\right]\right]$. □

A direct corollary of Lemma 7 is that $\sum_i \tau_i$ can be bounded using PREV.

Lemma 8. $\sum_{i\in[n]} \tau_i \leq 8\cdot \text{PREV}$.

Proof. By Constraint (6) of the LP in Figure 3 (or Figure 4), for every item j,

$$\sum_{i\in[n]} \Pr_{\substack{t_{ij}\sim D_{ij}\\ \beta_{ij}\sim\mathcal{B}_{ij}}}[V_{ij}(t_{ij}) \geq \max(\beta_{ij}, Q_j+\tau_i)] \leq \sum_{i\in[n]} \Pr_{\substack{t_{ij}\sim D_{ij}\\ \beta_{ij}\sim\mathcal{B}_{ij}}}[V_{ij}(t_{ij}) \geq \beta_{ij}]$$
$$= \sum_{i\in[n]}\sum_{\beta_{ij}\in\mathcal{V}_{ij}} \hat{\lambda}_{ij}(\beta_{ij}) \Pr_{t_{ij}}[V_{ij}(t_{ij}) \geq \beta_{ij}] \leq \frac{1}{2}$$

By the definition of τ_i, for every i we have

$$\sum_{j\in[m]} \Pr_{t_{ij}\sim D_{ij},\beta_{ij}\sim\mathcal{B}_{ij}}[V_{ij}(t_{ij}) \geq \max(\beta_{ij}, Q_j+\tau_i)]] \leq \frac{1}{2}.$$

Thus by Lemma 7,

$$4\cdot \text{PREV} \geq \sum_{i\in[n],j\in[m]} \mathop{\mathbb{E}}_{\beta_{ij}\sim\mathcal{B}_{ij}}\left[\max(\beta_{ij}, Q_j+\tau_i)\cdot \Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij}) \geq \max(\beta_{ij}, Q_j+\tau_i)]\right]$$
$$\geq \sum_{i\in[n]} \tau_i \sum_{j\in[m]} \mathop{\mathbb{E}}_{\beta_{ij}\sim\mathcal{B}_{ij}}\left[\Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij}) \geq \max(\beta_{ij}, Q_j+\tau_i)]\right]$$
$$= \frac{1}{2}\sum_{i\in[n]} \tau_i,$$

The last equality comes from the fact that by the definition of τ_i, $\sum_j \Pr_{t_{ij}\sim D_{ij},\beta_{ij}\sim\mathcal{B}_{ij}}[V_{ij}(t_{ij}) \geq \max(\beta_{ij}, Q_j+\tau_i)]] = \frac{1}{2}$ for all i such that $\tau_i > 0$ (see Footnote 21). We finish the proof. $\square$

Lemma 9. *[(Restatement of Lemma 6)] For every $j \in [m]$, $Q_j \geq \hat{Q}_j$. Moreover,*

$$\sum_{j\in[m]} Q_j \leq \sum_{j\in[m]} \hat{Q}_j + 236.5 \cdot \text{PREV}$$

Proof. For every j, it's clear that $\hat{Q}_j \leq Q_j$ by the definition of $\hat{Q}_j$. It remains to bound $\sum_j(Q_j - \hat{Q}_j)$. In the proof, we abuse the notation and let $\hat{\lambda}_{ij}(\beta_{ij}) = \sum_{\delta_{ij}\in\Delta} \hat{\lambda}_{ij}(\beta_{ij}, \delta_{ij})$. Also for every $i, j, t_{ij}, \beta_{ij} \in \mathcal{V}_{ij}$, let $\lambda_{ij}(t_{ij}, \beta_{ij}) = \sum_{\delta_{ij}\in\Delta} \lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij})$.

$$\begin{aligned}
&2\sum_{j\in[m]}\left(Q_j - \hat{Q}_j\right) \\
\leq &\sum_j\sum_i\sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij})V_{ij}(t_{ij}) \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta}} \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \beta_{ij}+\delta_{ij}] \\
\leq &\sum_j\sum_i\sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij}) \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta}} \left(\beta_{ij} + (V_{ij}(t_{ij})-\beta_{ij})^+\right)\cdot\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \beta_{ij}+\delta_{ij}] \\
\leq &\sum_j\sum_i\sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij}) \sum_{\beta_{ij}\in\mathcal{V}_{ij}} \beta_{ij}\cdot\lambda_{ij}(t_{ij},\beta_{ij}) \\
&+\sum_j\sum_i\sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij}) \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta}} (V_{ij}(t_{ij})-\beta_{ij})^+\cdot\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \beta_{ij}+\delta_{ij}]
\end{aligned}$$

Here the first inequality uses the fact that $\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\cdot\mathbb{1}[V_{ij}(t_{ij}) \leq \beta_{ij}+\delta_{ij}]$ and $\lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij})\mathbb{1}[V_{ij}(t_{ij}) \leq \min\{\beta_{ij}+\delta_{ij}, Q_j+\tau_i\}]$ can differ only when $V_{ij}(t_{ij}) > Q_j+\tau_i \wedge V_{ij}(t_{ij}) \leq \beta_{ij}+\delta_{ij}$. In the last inequality, we drop the indicator $\mathbb{1}[V_{ij}(t_{ij}) \leq \beta_{ij}+\delta_{ij}]$ for the first term.

We bound the first term:

$$
\begin{aligned}
&\sum_j \sum_i \sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij}) \cdot \sum_{\beta_{ij}\in\mathcal{V}_{ij}} \beta_{ij}\cdot\lambda_{ij}(t_{ij},\beta_{ij}) \\
=&\sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \beta_{ij}<Q_j+\tau_i}} \beta_{ij} \sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij})\cdot\lambda_{ij}(t_{ij},\beta_{ij}) + \sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \beta_{ij}\geq Q_j+\tau_i}} \beta_{ij} \sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij})\cdot\lambda_{ij}(t_{ij},\beta_{ij}) \\
\leq&\sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \beta_{ij}<Q_j+\tau_i}} \beta_{ij} \sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij})\cdot\lambda_{ij}(t_{ij},\beta_{ij}) \\
&\qquad + \sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}^0_{ij}\\ \beta_{ij}\geq Q_j+\tau_i}} \beta^+_{ij} \sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij})\cdot\Big(\lambda_{ij}(t_{ij},\beta_{ij})+\lambda_{ij}(t_{ij},\beta^+_{ij})\Big) \\
\leq&\sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \beta_{ij}<Q_j+\tau_i}} \hat{\lambda}_{ij}(\beta_{ij})\cdot\beta_{ij} \sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij}) + \sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}^0_{ij}\\ \beta_{ij}\geq Q_j+\tau_i}} 2\hat{\lambda}_{ij}(\beta_{ij})\cdot\beta^+_{ij}\cdot\Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij})\geq\beta_{ij}] \\
&\qquad + \sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}^+_{ij}\\ \beta_{ij}\geq Q_j+\tau_i}} 2\hat{\lambda}_{ij}(\beta_{ij})\cdot\beta_{ij}\cdot\Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij})\geq\beta_{ij}] \\
\leq&\sum_{i,j} \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \beta_{ij}<Q_j+\tau_i}} \hat{\lambda}_{ij}(\beta_{ij})\cdot\beta_{ij} \sum_{t_{ij}:V_{ij}(t_{ij})\geq Q_j+\tau_i} f_{ij}(t_{ij}) + \sum_{i,j}\left(\sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \beta_{ij}\geq Q_j+\tau_i}} 2\hat{\lambda}_{ij}(\beta_{ij})\cdot\beta_{ij}\cdot\Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij})\geq\beta_{ij}] + \varepsilon_r|\mathcal{V}^0_{ij}|\right) \\
\leq&2\sum_{i,j}\sum_{\beta_{ij}\in\mathcal{V}_{ij}} \hat{\lambda}_{ij}(\beta_{ij})\cdot\beta_{ij}\cdot\Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij})\geq\max(\beta_{ij},Q_j+\tau_i)] + \varepsilon_r\cdot\sum_{i,j}|\mathcal{V}^0_{ij}| \\
=&2\sum_{i,j}\mathbb{E}_{\beta_{ij}\sim\mathcal{B}_{ij}}\left[\beta_{ij}\cdot\Pr_{t_{ij}\sim D_{ij}}[V_{ij}(t_{ij})\geq\max(\beta_{ij},Q_j+\tau_i)]\right] + \varepsilon_r\cdot\sum_{i,j}|\mathcal{V}^0_{ij}| \\
\leq&8\cdot\text{PREV} + \varepsilon_r\cdot\sum_{i,j}|\mathcal{V}^0_{ij}| \\
\leq&9\cdot\text{PREV}
\end{aligned} \tag{4}
$$

The first inequality comes from the fact that for $\beta_{ij}\in\mathcal{V}^0_{ij}$, then $\beta_{ij}<\beta^+_{ij}$ and from the fact that for sufficiently small $\varepsilon_r>0$, then $Q_j+\tau_i\leq\beta_{ij}$ iff $Q_j+\tau_i\leq\beta^+_{ij}$. The second inequality comes from Constraint (4) and (7) of the LP in Figure 3 (or Figure 4).[22] For the second last inequality, notice that by Constraint (6) of the LP in Figure 3 (or Figure 4), for every item j,

$$
\begin{aligned}
\sum_i \Pr_{\substack{t_{ij}\sim D_{ij}\\ \beta_{ij}\sim\mathcal{B}_{ij}}}[V_{ij}(t_{ij})\geq\max(\beta_{ij},Q_j+\tau_i)]] &\leq \sum_i \Pr_{\substack{t_{ij}\sim D_{ij}\\ \beta_{ij}\sim\mathcal{B}_{ij}}}[V_{ij}(t_{ij})\geq\beta_{ij}] \\
&= \sum_i \sum_{\beta_{ij}\in\mathcal{V}_{ij}} \hat{\lambda}_{ij}(\beta_{ij})\Pr_{t_{ij}}[V_{ij}(t_{ij})\geq\beta_{ij}] \leq \frac{1}{2}
\end{aligned}
$$

By the definition of τ_i, for every buyer i we have

$$\sum_j \Pr_{t_{ij}\sim D_{ij},\beta_{ij}\sim\mathcal{B}_{ij}}[V_{ij}(t_{ij})\geq\max(\beta_{ij},Q_j+\tau_i)]] \leq \frac{1}{2}$$

[22] Note that this is the only place that Constraint (7) is used in our proof.

The second last inequality then follows from Lemma 7. The last inequality follows from the fact that $\varepsilon_r \le \frac{\textsc{PRev}}{\sum_{i,j}|\mathcal{V}_{ij}^0|}$. For the second term, we have

$$\sum_j \sum_i \sum_{t_{ij}:V_{ij}(t_{ij})\ge Q_j+\tau_i} f_{ij}(t_{ij}) \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta_i}} (V_{ij}(t_{ij})-\beta_{ij})^+ \cdot \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \le \beta_{ij}+\delta_{ij}]$$

$$\le \sum_j \sum_i \sum_{\substack{\beta_{ij}\in\mathcal{V}_{ij}\\ \delta_{ij}\in\Delta_i}} \hat{\lambda}_{ij}(\beta_{ij},\delta_{ij}) \cdot \sum_{t_{ij}} f_{ij}(t_{ij})(V_{ij}(t_{ij})-\beta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \le \beta_{ij}+\delta_{ij} \wedge V_{ij}(t_{ij}) \ge \max(\beta_{ij}, Q_j+\tau_i)]$$

$$= \sum_{i,j} \mathbb{E}_{(\beta_{ij},\delta_{ij})\sim\mathcal{C}_{ij}} \left[\sum_{t_{ij}} f_{ij}(t_{ij}) u_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right],$$

where

$$u_{ij}(t_{ij},\beta_{ij},\delta_{ij}) = (V_{ij}(t_{ij})-\beta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \le \beta_{ij}+\delta_{ij} \wedge V_{ij}(t_{ij}) \ge \max(\beta_{ij}, Q_j+\tau_i)]$$

We notice that by the definition of τ_i, the following inequality holds for every buyer i.

$$\sum_j \Pr_{t_{ij}\sim D_{ij},(\beta_{ij},\delta_{ij})\sim\mathcal{C}_{ij}}[u_{ij}(t_{ij},\beta_{ij},\delta_{ij})>0] \le \sum_j \Pr_{t_{ij}\sim D_{ij},\beta_{ij}\sim\mathcal{B}_{ij}}[V_{ij}(t_{ij}) \ge \max(\beta_{ij},Q_j+\tau_i)]] \le \frac{1}{2} \quad (5)$$

Denote $\mathcal{C}_i = \times_{j=1}^m \mathcal{C}_{ij}$ and $\beta_i = (\beta_{ij})_{j\in[m]}, \delta_i = (\delta_{ij})_{j\in[m]}$, we have

$$\sum_i \mathbb{E}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\sum_{t_i} f_i(t_i) \cdot \max_j u_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right]$$

$$\ge \sum_i \mathbb{E}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\sum_j \sum_{t_{ij}} f_{ij}(t_{ij}) \cdot u_{ij}(t_{ij},\beta_{ij},\delta_{ij}) \cdot \prod_{k\ne j} \Pr_{t_{ik}\sim D_{ik}}[u_{ik}(t_{ik},\beta_{ik},\delta_{ik})=0]\right]$$

$$= \sum_i \sum_j \mathbb{E}_{(\beta_{ij},\delta_{ij})\sim\mathcal{C}_{ij}}\left[\sum_{t_{ij}} f_{ij}(t_{ij}) \cdot u_{ij}(t_{ij},\beta_{ij},\delta_{ij}) \cdot \prod_{k\ne j} \Pr_{t_{ik}\sim D_{ik},(\beta_{ik},\delta_{ik})\sim\mathcal{C}_{ik}}[u_{ik}(t_{ik},\beta_{ik},\delta_{ik})=0]\right]$$

$$\ge \frac{1}{2} \cdot \sum_{i,j} \mathbb{E}_{(\beta_{ij},\delta_{ij})\sim\mathcal{C}_{ij}}\left[\sum_{t_{ij}} f_{ij}(t_{ij}) u_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right]$$

Here the equality uses the fact that all $\mathcal{C}_{ij}$'s are independent. The last inequality comes from Inequality (5) and the union bound. Now the second term is bounded by

$$2 \cdot \sum_i \mathbb{E}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\sum_{t_i} f_i(t_i) \cdot \max_j u_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right]$$

$$\le 2 \cdot \sum_i \mathbb{E}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\sum_{t_i} f_i(t_i) \cdot \max_{j\in[m]} \left\{(V_{ij}(t_{ij})-\beta_{ij})^+ \cdot \mathbb{1}[V_{ij}(t_{ij}) \le \beta_{ij}+\delta_{ij}]\right\}\right]$$

Definition 17. *For every $i, j, t_i \in \mathcal{T}_i, S \subseteq [m]$, let*

$$\eta_i(t_i, S) = \mathbb{E}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S}\left\{(V_{ij}(t_{ij})-\beta_{ij})^+ \cdot \mathbb{1}[V_{ij}(t_{ij}) \le \beta_{ij}+\delta_{ij}]\right\}\right]$$

Therefore, we can rewrite $2\cdot\sum_i \mathbb{E}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\sum_{t_i} f_i(t_i)\cdot \max_{j\in[m]}\left\{(V_{ij}(t_{ij})-\beta_{ij})^+\cdot \mathbb{1}[V_{ij}(t_{ij})\leq \beta_{ij}+\delta_{ij}]\right\}\right]$ as $\sum_i\sum_{t_i} f_i(t_i)\cdot\eta(t_i,[m])$.

Definition 18. *A function $v(\cdot,\cdot)$ is a-**Lipschitz** if for any type $t,t'\in T$, and set $X,Y\subseteq[m]$,*

$$\left|v(t,X)-v(t',Y)\right|\leq a\cdot\left(|X\Delta Y|+\left|\{j\in X\cap Y: t_j\neq t'_j\}\right|\right),$$

where $X\Delta Y=\big(X\setminus Y\big)\cup\big(Y\setminus X\big)$ is the symmetric difference between X and Y.

Lemma 10. *For every i, $\eta_i(\cdot,\cdot)$ is subadditive over independent items and d_i-Lipschitz. Note that d_i is the variable in the LP in Figure 3.*

Proof. For every $i,j,t_{ij},\beta_{ij},\delta_{ij}$, denote $h_{ij}(t_{ij},\beta_{ij},\delta_{ij})=(V_{ij}(t_{ij})-\beta_{ij})^+\cdot\mathbb{1}[V_{ij}(t_{ij})\leq\beta_{ij}+\delta_{ij}]$.

We will first verify that for each $t_i\in T_i$, $\eta_i(t_i,\cdot)$ is monotone, subadditive and has no externalities.

Monotonicity: Let $S_1\subseteq S_2\subseteq[m]$. Then:

$$\eta_i(t_i,S_1)=\mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S_1}\left\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right\}\right]\leq\mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S_2}\left\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right\}\right]=\eta_i(t_i,S_2)$$

Subadditivity: For any set $S_1,S_2,S_3\subseteq[m]$ such that $S_1\cup S_2=S_3$, it holds that:

$$\begin{aligned}\eta_i(t_i,S_3)&=\mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S_3}\left\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right\}\right]\\&\leq\mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S_1}\left\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right\}+\max_{j\in S_2}\left\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right\}\right]\\&=\eta_i(t_i,S_1)+\eta_i(t_i,S_2)\end{aligned}$$

The first inequality follows because $S_1\cup S_2=S_3$.

No externalities: for every $S\subseteq[m]$ and $t_i,t'_i\in\mathcal{T}_i$ such that $t_{ij}=t'_{ij}$ for every $j\in S$, we have

$$\eta_i(t_i,S)=\mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S}\left\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right\}\right]=\mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S}\left\{h_{ij}(t'_{ij},\beta_{ij},\delta_{ij})\right\}\right]=\eta_i(t'_i,S)$$

Now we are going to prove that $\eta_i(\cdot,\cdot)$ is d_i-Lipschitz. For $t_i,t'_i\in\mathcal{T}_i$ and $X,Y\subseteq[m]$, let $Z=\{j\in X\cap Y\wedge t_{ij}=t'_{ij}\}$. It is enough to show that:

$$\begin{aligned}\eta_i(t_i,X)-\eta_i(t'_i,Y)\leq\big(|X\setminus Y|+(|X\cap Y|-|Z|)\big)\cdot d_i=(|X|-|Z|)\cdot d_i\\\eta_i(t'_i,Y)-\eta_i(t_i,X)\leq\big(|Y\setminus X|+(|X\cap Y|-|Z|)\big)\cdot d_i=(|Y|-|Z|)\cdot d_i\end{aligned}$$

We are only going to show $\eta_i(t_i,X)-\eta_i(t'_i,Y)\leq(|X|-|Z|)\cdot d_i$, since the other case is similar. Because $\eta_i(t_i,\cdot)$ is monotone, it suffices to prove that:

$$\eta_i(t_i,X)-\eta_i(t'_i,Y)\leq\eta_i(t_i,X)-\eta_i(t'_i,Z)\leq(|X|-|Z|)\cdot d_i$$

For each $j \in Z$, $t_{ij} = t'_{ij}$, which implies that $\eta_i(t'_i, Z) = \eta_i(t_i, Z)$. Note that:

$$\begin{aligned}
&\eta_i(t_i, X) - \eta_i(t'_i, Z) \\
=&\eta_i(t_i, X) - \eta_i(t_i, Z) \\
=& \mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in X}\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\} - \max_{j\in Z}\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\}\right] \\
\leq& \mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in Z}\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\} + \sum_{j\in X\backslash Z} h_{ij}(t_{ij},\beta_{ij},\delta_{ij}) - \max_{j\in Z}\{h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\}\right] \\
=& \mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\sum_{j\in X\backslash Z} h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right] \\
=& \sum_{j\in X\backslash Z}\mathop{\mathbb{E}}_{(\beta_{ij},\delta_{ij})\sim\mathcal{C}_{ij}}\left[h_{ij}(t_{ij},\beta_{ij},\delta_{ij})\right] \\
\leq& \sum_{j\in X\backslash Z}\mathop{\mathbb{E}}_{(\beta_{ij},\delta_{ij})\sim\mathcal{C}_{ij}}\left[\delta_{ij}\right] \\
\leq&(|X| - |Z|)d_i
\end{aligned}$$

Here the second last inequality follows by Constraint (8) of the LP in Figure 3 (or Figure 4). ☐

Lemma 11. *[CZ17] Let $g(t,\cdot)$ with $t \sim D = \prod_j D_j$ be a function drawn from a distribution that is subadditive over independent items of ground set I. If $g(\cdot,\cdot)$ is c-Lipschitz, then if we let a be the median of the value of the grand bundle $g(t, I)$, i.e. $a = \inf\left\{x \geq 0 : \Pr_t[g(t, I) \leq x] \geq \frac{1}{2}\right\}$,*

$$\mathbb{E}_t[g(t, I)] \leq 2a + \frac{5c}{2}.$$

To finish the proof of Lemma 9, we will bound $\sum_i \sum_{t_i} f_i(t_i) \cdot \eta_i(t_i, [m])$ using a modified two-part tariff mechanism. Consider the following variant of two-part tariff $\mathcal{M}$, with a randomized posted price $\beta_{ij} \sim \mathcal{B}_{ij}$ for buyer i and item j, and restricting the buyer to purchase at most one item. The procedure of the mechanism is shown in Mechanism 2 [23].

Mechanism 2 The Rationed Two-part Tariff with Randomized Posted Price $\mathcal{M}$

0: Before the mechanism starts, the seller determines a **distribution** of posted price $\mathcal{B}_{ij}$ for every buyer i and item j. Recall that $\mathcal{B}_{ij}$ is the marginal distribution of β_{ij}.

1: Bidders arrive sequentially in the lexicographical order.

2: When every buyer i arrives, the seller shows her the set of available items $S_i(t_{<i}, \beta_{<i}) \subseteq [m]$ (see the remark below), as well as the *distribution* of the posted prices $\{\mathcal{B}_{ij}\}_{j\in[m]}$.

3: Buyer i is asked to pay an *entry fee* defined as follows:

$$\xi_i(S_i(t_{<i}, \beta_{<i})) = \text{MEDIAN}_{t_i\sim D_i}\{\eta_i(t_i, S_i(t_{<i}, \beta_{<i}))\}.$$

Here $\text{MEDIAN}_x[h(x)]$ denotes the median of a non-negative function $h(x)$ on random variable x, i.e. $\text{MEDIAN}_x[h(x)] = \inf\{a \geq 0 : \Pr_x[h(x) \leq a] \geq \frac{1}{2}\}$.

4: If buyer i (with type t_i) agrees to pay the entry fee, then the seller will sample a realized posted price $\beta_{ij} \sim \mathcal{B}_{ij}$ for every available item $j \in S_i(t_{<i}, \beta_{<i})$. The buyer is restricted to purchase **at most one** item. The buyer then either chooses her favorite item $j^* = \arg\max_{j\in S_i(t_{<i},\beta_{<i})}\left(V_{ij}(t_{ij}) - \beta_{ij}\right)$, and pays β_{ij^*}, or leaves with nothing if $V_{ij}(t_{ij}) < \beta_{ij}, \forall j \in S_i(t_{<i}, \beta_{<i})$. If the buyer refuses to pay the entry fee, she gets nothing and pays 0.

[23] The result holds for any buyers' order, we choose the lexicographical order to keep the notation light.

Remark 3. *We notice that in Mechanism 2, the set of available items when each buyer i comes to the auction, depends on both $t_{<i}$ and the realized prices for the first $i-1$ buyers (denoted by $\beta_{<i}$). Let $S_i(t_{<i},\beta_{<i})$ be the random set of available items when buyer i comes to the auction. Let $S_1(t_{<1},\beta_{<1})=[m]$.*

It's not hard to see that $\mathcal{M}$ stated in Mechanism 2 is BIC and IR: When every buyer sees the set of available items and the distribution of posted prices, she can calculate her expected surplus of this set, over the randomness of the posted prices. Then she will accept the entry fee if and only if the expected surplus is at least the entry fee.

In Mechanism 2, by union bound, for every item j,

$$\Pr[j\in S_i(t_{<i},\beta_{<i})]\geq 1-\sum_{k<i}\Pr_{t_{ij}\sim D_{ij},\beta_{ij}\sim\mathcal{B}_{ij}}[V_{ij}(t_{ij})\geq\beta_{ij}]\geq\frac{1}{2}\qquad\text{(Due to Constraint (6))}$$

We notice that for every realization of $t_{<i},\beta_{<i}$, after seeing the remaining item set $S_i(t_{<i},\beta_{<i})$, buyer i's expected surplus if she enters the mechanism is:

$$\mathop{\mathbb{E}}_{(\beta_i,\delta_i)\sim\mathcal{C}_i}\left[\max_{j\in S_i(t_{<i},\beta_{<i})}\left\{(V_{ij}(t_{ij})-\beta_{ij})^+\right\}\right]\geq\eta_i(t_i,S_i(t_{<i},\beta_{<i})).$$

Thus the buyer will accepts the entry fee with probability at least $1/2$. Hence

$$\begin{aligned}\textsc{Rev}(\mathcal{M})&\geq\frac{1}{2}\sum_i\mathop{\mathbb{E}}_{t_{<i},\beta_{<i}}[\xi_i(S_i(t_{<i},\beta_{<i}))]\geq\frac{1}{2}\cdot\sum_i(\frac{1}{2}\mathop{\mathbb{E}}_{t_i,t_{<i},\beta_{<i}}[\eta_i(t_i,S_i(t_{<i},\beta_{<i}))]-\frac{5}{4}d_i)\\&\geq\frac{1}{2}\cdot\sum_i(\frac{1}{4}\mathop{\mathbb{E}}_{t_i}[\eta_i(t_i,[m])]-\frac{5}{4}d_i)\geq\frac{1}{8}\cdot\sum_i\mathop{\mathbb{E}}_{t_i}[\eta_i(t_i,[m])]-5\cdot\text{P}\textsc{Rev}\end{aligned}\tag{6}$$

Here the second inequality is obtained by applying Lemma 11 to function $\eta_i(t_i,\cdot)$ and ground set $I=S_i(t_{<i},\beta_{<i})$ for every $t_{<i},\beta_{<i}$. The last inequality comes from constraint (9) of the LP in Figure 3. The third inequality comes from the following:

Fix any t_i,β_i,δ_i. Let $j^*=\arg\max_{j\in[m]}(V_{ij}(t_{ij})-\beta_{ij})^+\cdot\mathbb{1}[V_{ij}(t_{ij})\leq\beta_{ij}+\delta_{ij}]$. Then

$$\begin{aligned}&\mathop{\mathbb{E}}_{t_{<i},\beta_{<i}}\left[\max_{j\in S_i(t_{<i},\beta_{<i})}\left\{(V_{ij}(t_{ij})-\beta_{ij})^+\cdot\mathbb{1}[V_{ij}(t_{ij})\leq\beta_{ij}+\delta_{ij}]\right\}\right]\\\geq&\mathop{\mathbb{E}}_{t_{<i},\beta_{<i}}\left[\max_{j\in[m]}\left\{(V_{ij}(t_{ij})-\beta_{ij})^+\cdot\mathbb{1}[V_{ij}(t_{ij})\leq\beta_{ij}+\delta_{ij}]\right\}\cdot\mathbb{1}[j^*\in S_i(t_{<i},\beta_{<i})]\right]\\=&\max_{j\in[m]}\left\{(V_{ij}(t_{ij})-\beta_{ij})^+\cdot\mathbb{1}[V_{ij}(t_{ij})\leq\beta_{ij}+\delta_{ij}]\right\}\cdot\Pr_{t_{<i},\beta_{<i}}[j^*\in S_i(t_{<i},\beta_{<i})]\\\geq&\frac{1}{2}\cdot\max_{j\in[m]}\left\{(V_{ij}(t_{ij})-\beta_{ij})^+\cdot\mathbb{1}[V_{ij}(t_{ij})\leq\beta_{ij}+\delta_{ij}]\right\}\end{aligned}$$

Taking expectation over t_i and $(\beta_i,\delta_i)\sim\mathcal{C}_i$ on both sides, we have

$$\mathop{\mathbb{E}}_{t_i,t_{<i},\beta_{<i}}[\eta_i(t_i,S_i(t_{<i},\beta_{<i}))]\geq\frac{1}{2}\mathop{\mathbb{E}}_{t_i}[\eta_i(t_i,[m])],$$

which is exactly the third inequality of (6). By combining Inequalities (4) and (6), we have

$$2\sum_j(Q_j-\hat{Q}_j)\leq 89\cdot\text{P}\textsc{Rev}+16\cdot\textsc{Rev}(\mathcal{M})$$

Finally, since in $\mathcal{M}$, each buyer is restricted to purchase at most one item, it can be treated as a BIC and IR mechanism in the unit-demand setting. By [CHMS10b, KW12, CDW16], $\text{REV}(\mathcal{M}) \leq 24 \cdot \text{PREV}$. We complete our proof for Lemma 9.

□

C.6 Analyzing the Revenue of $\mathcal{M}_{\text{TPT}}$ and the Proof of Theorem 6

In this section, we will show that $\sum_{j\in[m]} Q_j$ can be bounded using the revenue of the two-part tariff $\mathcal{M}_{\text{TPT}}$ defined in Mechanism 1. To analyze the revenue of $\mathcal{M}_{\text{TPT}}$, we require the following definition.

Definition 19. *For any buyer i and type t_i, let $C_i(t_i) = \{j \in [m] \mid V_{ij}(t_{ij}) \leq Q_j + \tau_i\}$. For any $t_i \in \mathcal{T}_i$ and set $S \subseteq [m]$, let*

$$\mu_i(t_i, S) = \max_{S' \subseteq S} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right)$$

Lemma 12. *For every i, if $v_i(\cdot,\cdot)$ is subadditive over independent itmes, then $\mu_i(\cdot,\cdot)$ is subadditive over independent items and τ_i-Lipschitz.*

Proof. First we are going to show that $\mu_i(t,\cdot)$ is monotone. Note that for sets $S_1 \subseteq S_2$ the following holds:

$$\begin{aligned}
\mu_i(t_i, S_1) = &\max_{S' \subseteq S_1} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) \\
\leq &\max_{S' \subseteq S_2} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) \\
= &\mu_i(t_i, S_2)
\end{aligned}$$

Now we are going to show that $\mu_i(t,\cdot)$ is subadditive. Let $S_1, S_2, S_3 \subseteq [m]$ such that $S_1 \cup S_2 = S_3$ and $S_1 \cap S_2 = S_c$. Let $S_a = S_1 \backslash S_c$ and $S_b = S_2$, then for any $t_i \in T_i$ we have the following:

$$\begin{aligned}
\mu_i(t_i, S_3) = &\max_{S' \subseteq S_3} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) \\
\leq &\max_{S' \subseteq S_3} \left(v_i(t_i, \left(S_a \cap S'\right) \cap C_i(t_i)) + v_i(t_i, \left(S_b \cap S'\right) \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) \\
= &\max_{S' \subseteq S_3} \left(\left(v_i(t_i, \left(S_a \cap S'\right) \cap C_i(t_i)) - \sum_{j \in S_a \cap S'} Q_j \right) + \left(v_i(t_i, \left(S_b \cap S'\right) \cap C_i(t_i)) - \sum_{j \in S_b \cap S'} Q_j \right) \right) \\
\leq &\max_{S' \subseteq S_a} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) + \max_{S' \subseteq S_b} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) \\
= &\mu_i(t_i, S_a) + \mu_i(t_i, S_b) \\
\leq &\mu_i(t_i, S_1) + \mu_i(t_i, S_2)
\end{aligned}$$

The first inequality follows because $v_i(t_i,\cdot)$ is a subadditive function and $S_a \cup S_b = S_3$. The second inequality follows because $\max$ is subadditive. The final inequality follows from the fact that $S_b = S_2$, $S_1 \supseteq S_a$ and that $\mu_i(t_i,\cdot)$ is monotone.

We now prove that $\mu_i(\cdot,\cdot)$ has no externalities. Fix any $S \subseteq [m]$ and $t_i, t_i' \in \mathcal{T}_i$ such that $t_{ij} = t_{ij}'$ for all $j \in S$. We notice that by the definition of C_i, $S' \cap C_i(t_i) = S' \cap C_i(t_i')$ for all $S' \subseteq S$. Since $v_i(\cdot,\cdot)$ has no externalities, $v_i(t_i, S' \cap C_i(t_i)) = v_i(t_i', S' \cap C_i(t_i'))$ for every $S' \subseteq S$. Thus $\mu_i(t_i, S) = \mu_i(t_i', S)$.

Now we are going to show that $\mu_i(\cdot,\cdot)$ is τ_i-Lipschitz. Let $t_i, t_i' \in \mathcal{T}_i$ and $X, Y \subseteq [m]$ and $c = |\{j \in [m] : j \in X \Delta Y \text{ or } t_{ij} \neq t_{ij}'\}|$, [24] we need to show that:

$$|\mu_i(t_i, X) - \mu_i(t_i', Y)| \leq c \cdot \tau_i$$

Let $Z = \{j : j \in X \cap Y \text{ and } t_{ij} = t_{ij}'\}$. Since $\mu_i(t_i, \cdot)$ is monotone, in order to show that $\mu_i(\cdot,\cdot)$ is τ_i-Lipschitz, it is enough to show that

$$\mu_i(t_i, X) - \mu_i(t_i', Y) \leq \mu_i(t_i, X) - \mu_i(t_i', Z) \leq c \cdot \tau_i$$
$$\mu_i(t_i', Y) - \mu_i(t_i, X) \leq \mu_i(t_i', Y) - \mu_i(t_i, Z) \leq c \cdot \tau_i$$

We are only going to prove that $\mu_i(t_i, X) - \mu_i(t_i', Z) \leq c \cdot \tau_i$ since the other case is similar. Because for each $j \in Z$, $t_{i,j}' = t_{i,j}$, then $\mu_i(t_i', Z) = \mu_i(t_i, Z)$. We have

$$\begin{aligned}
\mu_i(t_i, X) &= \max_{S' \subseteq X} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) \\
&\leq \max_{S' \subseteq X} \left(\sum_{j \in S' \setminus Z} \big(v_i(t_i, \{j\} \cap C_i(t_i)) - Q_j\big) + \left(v_i(t_i, (Z \cap S') \cap C_i(t_i)) - \sum_{j \in Z \cap S'} Q_j \right) \right) \\
&\leq \max_{S' \subseteq X} \left(\sum_{j \in S' \setminus Z} \big(V_{ij}(t_{ij}) - Q_j\big)^+ \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] + \left(v_i(t_i, (Z \cap S') \cap C_i(t_i)) - \sum_{j \in Z \cap S'} Q_j \right) \right) \\
&= \max_{S' \subseteq Z} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) + \sum_{j \in X \setminus Z} \big(V_{ij}(t_{ij}) - Q_j\big)^+ \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] \\
&\leq \max_{S' \subseteq Z} \left(v_i(t_i, S' \cap C_i(t_i)) - \sum_{j \in S'} Q_j \right) + (|X| - |Z|)\tau_i \\
&\leq \mu_i(t_i, Z) + c \cdot \tau_i
\end{aligned}$$

The first inequality follows because $v_i(t_i, \cdot)$ is subadditive. The second inequality follows because $v_i(t_i, \{j\} \cap C_i(t_i)) - Q_j \leq (V_{ij}(t_{ij}) - Q_j)^+ \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i]$. □

Lemma 13. *For every type profile $t \in \mathcal{T}$, let $\text{SOLD}(t)$ be the set of items sold in mechanism $\mathcal{M}_{TPT}$. Then*

$$\begin{aligned}
\mathbb{E}_t\left[\sum_{i \in [n]} \mu_i\left(t_i, S_i(t_{<i})\right)\right] &\geq \sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot (2\hat{Q}_j - Q_j) \\
&\geq \sum_j \Pr_t\left[j \notin \text{SOLD}(t)\right] \cdot Q_j - 473 \cdot \text{PRev}
\end{aligned}$$

Proof. By the definition of polytope W_i, for every buyer i and $t_i \in \mathcal{T}_i$, there exists an vector of non-negative numbers $\{\sigma_{iS}^{(k)}(t_i)\}_{S \subseteq [m], k \in [K]}$, such that $\sum_{S,k} \sigma_{iS}^{(k)}(t_i) \leq 1$ and

$$\pi_{ij}(t_{ij}) = f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S: j \in S} \sum_{k \in [K]} \sigma_{iS}^{(k)}(t_{ij}, t_{i,-j}) \quad (7)$$

[24] Δ stands for the symmetric difference between two sets.

$$w_{ij}(t_{ij}) \cdot V_{ij}(t_{ij}) \leq f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j} \in \mathcal{T}_{i,-j}} f_{i,-j}(t_{i,-j}) \sum_{S:j \in S} \sum_k \sigma_{iS}^{(k)}(t_{ij}, t_{i,-j}) \cdot \alpha_{ij}^{(k)}(t_{ij}) \tag{8}$$

We have

$$\begin{aligned}
&\mathbb{E}_t \left[\sum_i \mu_i \left(t_i, S_i(t_{<i})\right) \right] \\
\geq &\sum_i \mathbb{E}_{t_i, t_{-i}} \left[\sum_{S \subseteq [m]} \sum_k \sigma_{iS}^{(k)}(t_i) \cdot \mu_i \left(t_i, S_i(t_{<i}) \cap S\right) \right] \\
\geq &\sum_i \mathbb{E}_{t_i, t_{-i}} \left[\sum_{S,k} \sigma_{iS}^{(k)}(t_i) \cdot \sum_{j \in S} \mathbb{1}\left[j \in S_i(t_{<i})\right] \cdot \left(\alpha_{ij}^{(k)}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - Q_j\right)^+ \right] \\
= &\sum_i \mathbb{E}_{t_i} \left[\sum_{j \in [m]} \Pr_{t_{-i}}[j \in S_i(t_{<i})] \cdot \sum_{S:j \in S} \sum_k \sigma_{iS}^{(k)}(t_i) \cdot \left(\alpha_{ij}^{(k)}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - Q_j\right)^+ \right] \\
\geq &\sum_i \sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot \mathbb{E}_{t_i} \left[\sum_{S:j \in S} \sum_k \sigma_{iS}^{(k)}(t_i) \cdot \left(\alpha_{ij}^{(k)}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - Q_j\right)^+ \right] \\
\geq &\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot \sum_i \sum_{t_i} f_i(t_i) \cdot \sum_{S:j \in S} \sum_k \sigma_{iS}^{(k)}(t_i) \cdot \left(\alpha_{ij}^{(k)}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - Q_j\right) \\
= &\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot \sum_i \sum_{t_{ij}} f_{ij}(t_{ij}) \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S:j \in S} \sum_k \sigma_{iS}^{(k)}(t_i) \cdot (\alpha_{ij}^{(k)}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - Q_j) \\
\geq &\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot \sum_i \sum_{t_{ij}} w_{ij}(t_{ij}) V_{ij}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - \sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot Q_j
\end{aligned}$$

The first inequality is because $\mu_i(t_i, S)$ is monotone in set S for any i, t_i and $\sum_{S,k} \sigma_{iS}^{(k)}(t_i) \leq 1$. For any fixed i, t_i and set S, if we let S' be the set of items that are in $S \cap S_i(t_{<i})$ and satisfy that $\alpha_{ij}^{(k)}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - Q_j \geq 0$. Clearly $S' \subseteq C_i(t_i)$. Then

$$\begin{aligned}
\mu_i \left(t_i, S_i(t_{<i}) \cap S\right) &\geq v_i(t_i, S') - \sum_{j \in S'} Q_j \\
&= \max_{k' \in [K]} \sum_{j \in S'} \alpha_{ij}^{(k')}(t_{ij}) - \sum_{j \in S'} Q_j \geq \sum_{j \in S'} \left(\alpha_{ij}^{(k)}(t_{ij}) - Q_j\right) \\
&= \sum_{j \in S'} \left(\alpha_{ij}^{(k)}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq Q_j + \tau_i] - Q_j\right) \qquad (S' \subseteq C_i(t_i))
\end{aligned}$$

This inequality is exactly the second inequality above. The third inequality is because $\Pr_{t_{<i}}[j \in S_i(t_{<i})] \geq \Pr_t[j \notin \text{SOLD}(t)]$ for all j and i, as the LHS is the probability that the item is not sold after the seller has visited the first $i-1$ buyers and the RHS is the probability that the item remains unsold till the end of the mechanism $\mathcal{M}_{\text{TPT}}$. The last inequality follows from Inequality (8) and Constraint (2) of the LP in Figure 3 (or in Figure 4):

$$\sum_i \sum_{t_{ij}} f_{ij}(t_{ij}) \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S:j \in S} \sum_k \sigma_{iS}^{(k)}(t_i) = \sum_i \sum_{t_{ij}} \pi_{ij}(t_{ij}) \leq 1$$

Notice that by Definition 6 and Constraint (3) of the LP in Figure 3 (or in Figure 4), for every i, j, t_{ij},

$$\sum_{\beta_{ij},\delta_{ij}} \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij}) = w_{ij}(t_{ij})/f_{ij}(t_{ij})$$

Thus we have

$$\begin{aligned}
&\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot \sum_i \sum_{t_{ij}} w_{ij}(t_{ij}) V_{ij}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \le Q_j + \tau_i] - \sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot Q_j \\
=&\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot \sum_i \sum_{t_{ij}} f_{ij}(t_{ij}) V_{ij}(t_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \le Q_j + \tau_i] \sum_{\beta_{ij},\delta_{ij}} \lambda_{ij}(t_{ij},\beta_{ij},\delta_{ij}) \\
&\qquad - \sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot Q_j \\
\ge&\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot (2\hat{Q}_j - Q_j) \quad \text{(Definition 16)} \\
=&\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot Q_j - \sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot 2(Q_j - \hat{Q}_j) \\
\ge&\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot Q_j - \sum_j 2(Q_j - \hat{Q}_j) \quad \text{(Lemma 9, } Q_j - \hat{Q}_j \ge 0 \text{ for all } j) \\
\ge&\sum_j \Pr_t[j \notin \text{SOLD}(t)] \cdot Q_j - 473 \cdot \text{PREV} \quad \text{(Lemma 9)}
\end{aligned}$$

□

Now we give the proof of Theorem 6. Note that this is also the proof of Theorem 3.

Proof of Theorem 6:

For every $i, t_{<i}$, we apply Lemma 11 to function $\mu_i(t_i, \cdot)$ and ground set $S_i(t_{<i})$. By Definition 19, we have

$$\mathbb{E}_{t_i}[\mu_i(t_i, S_i(t_{<i}))] \le 2 \cdot \text{MEDIAN}_{t_i}(\mu_i(t_i, S_i(t_{<i}))) + \frac{5}{2} \cdot \tau_i \tag{9}$$

We now bound the revenue of $\mathcal{M}_{\text{TPT}}$. For every $i \in [n]$, $t_i \in \mathcal{T}_i$ and $S \subseteq [m]$, let $\mu_i'(t_i, S) = \max_{S' \subseteq S}(v_i(t_i, S') - \sum_{j \in S'} Q_j)$ which is at least as large as $\mu_i(t_i, S)$. Then the surplus of buyer i with true type $\hat{t}_i$, for the set $S_i(t_{<i})$ is $\mu_i'(\hat{t}_i, S_i(t_{<i}))$. By Mechanism 1, the entry fee $\xi_i(S_i(t_{<i}), t_i') = \mu_i'(t_i', S_i(t_{<i}))$ for every sampled type t_i'. Thus for every $t_{<i}$, we have

$$\Pr_{\hat{t}_i, t_i' \sim D_i}\left[\mu_i'(\hat{t}_i, S_i(t_{<i})) \ge \xi_i(S_i(t_{<i}), t_i') \ge \text{MEDIAN}_{t_i}(\mu_i(t_i, S_i(t_{<i})))\right] \ge \frac{1}{8}$$

In other words, for every $t_{<i}$, with probability at least $1/8$, the buyer accepts the entry fee, and the entry fee is at least $\text{MEDIAN}_{t_i}(\mu_i(t_i, S_i(t_{<i})))$. Thus the revenue of $\mathcal{M}_{\text{TPT}}$ from the entry fee is at least

$$\begin{aligned}
&\frac{1}{8} \sum_i \mathbb{E}_{t_{<i}}\left[\text{MEDIAN}_{t_i}(\mu_i(t_i, S_i(t_{<i})))\right] \\
\ge&\frac{1}{16} \sum_i \mathbb{E}_{t_i, t_{<i}}[\mu_i(t_i, S_i(t_{<i}))] - \frac{5}{32} \cdot \sum_i \tau_i \quad \text{(Inequality (9))} \\
\ge&\frac{1}{16} \sum_j \Pr_t\left[j \notin \text{SOLD}(t)\right] \cdot Q_j - \frac{493}{16} \cdot \text{PREV} \quad \text{(Lemma 13 and Lemma 8)}
\end{aligned}$$

We notice that for $\mathcal{M}_{\text{TPT}}$, the revenue from the posted prices are $\sum_j \Pr[j \in \text{SOLD}(t)] \cdot Q_j$. Thus

$$\begin{aligned}\textsc{Rev}(\mathcal{M}_{\text{TPT}}) \geq & \frac{1}{16}\sum_j \Pr_t\left[j \notin \text{SOLD}(t)\right] \cdot Q_j - \frac{493}{16} \cdot \text{PRev} + \sum_j \Pr[j \in \text{SOLD}(t)] \cdot Q_j \\ \geq & \frac{1}{16} \cdot \sum_{j\in[m]} Q_j - \frac{493}{16} \cdot \text{PRev}\end{aligned}$$

Thus

$$2 \cdot \sum_j Q_j \leq 986 \cdot \text{PRev} + 32 \cdot \textsc{Rev}(\mathcal{M}_{\text{TPT}})$$

We then show that $\mathcal{M}_{\text{TPT}}$ can be computed in polynomial time: By Definition 6, the posted price Q_j can be computed in time poly$(n, m, \sum_{i,j} |\mathcal{T}_{ij}|)$, given the feasible solution of the LP in Figure 3 (or in Figure 4). Given the set of available items $S_i(t_{<i})$, for every sampled type t_i', calculating the entry fee requires a single query from the demand oracle. For every buyer i with reported type t_i, the mechanism requires a single query from the demand oracle to obtain her favorite bundle among the set of available items, under prices $\{Q_j\}_{j\in[m]}$, and to determine whether the buyer will accept the entry fee.

Lastly, by Lemma 2, we can compute an RPP $\mathcal{M}_{\text{PP}}$ with the desired running time and query complexity, such that $\mathcal{M}_{\text{PP}} \geq \frac{1}{6.75}(1 - \frac{1}{nm}) \cdot \text{PRev}$. We finish our proof. □

D Multiplicative Approximation of Down-Monotone and Boxable Polytopes

In this section, we provide a proof of Theorem 2 and prove Theorem 1 for constrained-additive valuations using Theorem 2. We restate the theorem here.

Theorem 7. *(Restatement of Theorem 1 for constrained-additive valuations) Let $T = \sum_{i,j} |\mathcal{T}_{ij}|$ and b be the bit complexity of the problem instance (Definition 3). For constrained-additive buyers, for any $\delta > 0$, there exists an algorithm that computes a rationed posted price mechanism or a two-part tariff mechanism, such that the revenue of the mechanism is at least $c \cdot \text{OPT}$ for some absolute constant $c > 0$ with probability $1-\delta-\frac{2}{nm}$. Our algorithm assumes query access to a value oracle and a demand oracle of buyers' valuations, and has running time poly$(n, m, T, b, \log(1/\delta))$.*

For Theorem 2, we indeed prove a result for a natural family of polytopes. Throughout this section, we assume that the polytope we consider is *down-monotone*. Formally, a polytope $\mathcal{P} \subseteq [0, 1]^d$ is down-monotone if and only if for every $\boldsymbol{x} \in \mathcal{P}$ and $\boldsymbol{0} \leq \boldsymbol{x}' \leq \boldsymbol{x}$, we have $\boldsymbol{x}' \in \mathcal{P}$. To state our result, we need the following definitions.

Definition 20. *For any two sets $A, B \subseteq \mathbb{R}^d$, we denote by $A + B$ the* Minkowski addition *of set A and set B where:*

$$A + B = \{a + b : a \in A \text{ and } b \in B\}$$

Note that if both A and B are convex, then $A + B$ is also convex.

Definition 21. *Let $\mathcal{P}$ be a convex polytope, we define $a \cdot \mathcal{P} := \{a\boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\}$ for any $a \geq 0$.*

Definition 22. *Let ℓ be a finite integer. For any set of convex sets $\{\mathcal{P}_i\}_{i\in[\ell]}$ and a distribution $\mathcal{D} = \{q_i\}_{i\in[\ell]}$, the set $\mathcal{P} = \sum_{i\in[\ell]} q_i\mathcal{P}_i$ is called the mixture of $\{\mathcal{P}_i\}_{i\in[\ell]}$ over distribution $\mathcal{D}$.*

Definition 23. *Let $\mathcal{P}, \mathcal{Q} \subseteq [0, 1]^d$ be down-monotone polytopes. For each coordinate $j \in [d]$, we define the width of $\mathcal{P}$ at coordinate j as $l_j(\mathcal{P}) = \max_{\boldsymbol{x}\in\mathcal{P}} x_j$. For any $\varepsilon > 0$, we define the $(\varepsilon, \mathcal{Q})$-truncated polytope of $\mathcal{P}$ (denoted as $\mathcal{P}^{tr(\varepsilon,\mathcal{Q})}$) in the following way: $\boldsymbol{x} \in \mathcal{P}^{tr(\varepsilon,\mathcal{Q})}$ if and only if there exists $\boldsymbol{x}' \in \mathcal{P}$ such that $x_j = x_j' \cdot \mathbb{1}[l_j(\mathcal{Q}) \geq \varepsilon], \forall j \in [d]$. We notice that since $\mathcal{P}$ is down-monotone, $\mathcal{P}^{tr(\varepsilon,\mathcal{Q})} \subseteq \mathcal{P}$. Moreover, $\mathcal{P}^{tr(\varepsilon,\mathcal{Q})}$ is convex if $\mathcal{P}$ is convex. We also use $\mathcal{P}^{tr(\varepsilon)}$ to denote $\mathcal{P}^{tr(\varepsilon,\mathcal{P})}$.*

Definition 24. *Let $\mathcal{P} \subseteq [0,1]^d$. For any $\varepsilon > 0$, define the ε-box polytope $\mathcal{P}^{box(\varepsilon)}$ of $\mathcal{P}$ as follows: $\mathcal{P}^{box(\varepsilon)} = \{\boldsymbol{x} \subseteq [0,1]^d : x_j \leq \min\left(\varepsilon, l_j(\mathcal{P})\right), \forall j \in [d]\}$. Clearly, $\mathcal{P}^{box(\varepsilon)}$ is convex.*

Theorem 8 is the main theorem of this section. We prove that if $\mathcal{P}$ is a mixture of a set of down-monotone, convex polytopes $\{\mathcal{P}_i\}_{i\in[\ell]}$, and $\mathcal{P}$ contains the polytope $c \cdot \mathcal{P}^{box(\varepsilon)}$ for some $c \leq 1$, then there exists another down-monotone, convex polytope $\widehat{\mathcal{P}}$ sandwiched between $c/6 \cdot \mathcal{P}$ and $\mathcal{P}$. And more importantly, we have an efficient separation oracle for $\widehat{\mathcal{P}}$, whose running time is *independent of* ℓ, as long as we can efficiently optimize any linear objective for every $\mathcal{P}_i$. The key feature of our separation oracle for $\widehat{\mathcal{P}}$ is that its running time does not depend on ℓ, as in our applications, ℓ is usually exponential in the input size.

Theorem 8. *Let ℓ be a positive integer, and $\mathcal{P} \subseteq [0,1]^d$ be a mixture of $\{\mathcal{P}_i\}_{i\in[\ell]}$ over distribution $\mathcal{D} = \{q_i\}_{i\in[\ell]}$, where for each $i \in [\ell]$, $\mathcal{P}_i \subseteq [0,1]^d$ is a convex and down-monotone polytope. Suppose for every $i \in [\ell]$, there exists an oracle $\mathcal{Q}_i(\cdot)$, whose output $\mathcal{Q}_i(\boldsymbol{a}) \in \arg\max\{\boldsymbol{a}\cdot\boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_i\}$ for any input $\boldsymbol{a} \in \mathbb{R}^d$. Given $\{l_j(\mathcal{P})\}_{j\in[d]}$, suppose $c \cdot \mathcal{P}^{box(\varepsilon)} \subseteq \mathcal{P}$ for some $\varepsilon > 0$ and $c \in (0,1]$. Let b be an upper bound on the bit complexity of $\mathcal{Q}_i(\boldsymbol{a})$ for all $i \in [\ell]$ and $\boldsymbol{a} \in \mathbb{R}^d$, as well as on the bit complexity of $l_j(\mathcal{P})$ for all $j \in [d]$. Let the parameter $k \geq \Omega\left(d^4\left(b + \log\left(\frac{1}{\varepsilon}\right)\right)\right)$. We can construct a convex and down-monotone polytope $\widehat{\mathcal{P}}$ using $N = \left\lceil \frac{8kd}{\varepsilon^2} \right\rceil$ samples from $\mathcal{D}$ such that with probability at least $1 - 2de^{-2dk}$, the following guarantees hold:*

1. *$\frac{c}{6} \cdot \mathcal{P} \subseteq \widehat{\mathcal{P}} \subseteq \mathcal{P}$.*
2. *There exists a separation oracle SO for $\widehat{\mathcal{P}}$, whose running time on input with bit complexity b', is poly $\left(b, b', k, d, \frac{1}{\varepsilon}\right)$ and requires poly $\left(b, b', k, d, \frac{1}{\varepsilon}\right)$ queries to oracles in $\{\mathcal{Q}_i\}_{i\in[\ell]}$ with inputs of bit complexity at most poly $\left(b, b', k, d, \frac{1}{\varepsilon}\right)$.*

The complete proof of Theorem 8 is postponed to Appendix D.1. Here we give a sketch of the proof. We first prove that if the polytope $\mathcal{P}$ contains c times the ε-box polytope, then the convex set $\frac{c}{2}(\mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)})$ is sandwiched between $\frac{c}{2}\mathcal{P}$ and $\mathcal{P}$ (Lemma 14 in Appendix D.1). Next, we construct the polytope $\widehat{\mathcal{P}}$ that is close to $\frac{c}{2}(\mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)})$. For $\varepsilon > 0$ and every $i \in [\ell]$, let $\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}$ be the $(\varepsilon, \mathcal{P})$-truncated polytope of $\mathcal{P}_i$. It is clear that $\mathcal{P}^{tr(\varepsilon)}$ is a mixture of $\{\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}_{i\in[\ell]}$ over distribution $\mathcal{D}$. We construct our polytope $\widehat{\mathcal{P}}$ using $\widehat{\mathcal{P}}^{tr(\varepsilon)}$, the mixture of $\{\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}_{i\in[\ell]}$ over an empirical distribution $\widehat{\mathcal{D}}$ of $\mathcal{D}$. Cai et al. [CDW12b, CDW13a] proved that with polynomially many samples from $\mathcal{D}$, $\widehat{\mathcal{P}}^{tr(\varepsilon)}$ and $\mathcal{P}^{tr(\varepsilon)}$ are close within an additive error ε in the ℓ_∞-norm (Theorem 9 in Appendix D.1). By choosing $\widehat{\mathcal{P}} = \frac{c}{3}(\widehat{\mathcal{P}}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)})$, we show that $\widehat{\mathcal{P}}$ is a multiplicative approximation to $\mathcal{P}$.

To apply Theorem 8 to the single-bidder marginal reduced form polytope W_i, we first show that W_i is a mixture of a set of polytopes $\{W_{i,t_i}\}_{t_i\in\mathcal{T}_i}$ over D_i, where each W_{i,t_i} contains "all feasible single-bidder marginal reduced forms" for a specific type t_i (Definition 25 in Appendix D.2). For every t_i, we can maximize any linear objective over W_{i,t_i} via a query to the demand oracle. Finally, we prove that W_i contains (c times) the ε-box polytope of itself, for some $c \in (0,1)$ and $\varepsilon > 0$.

D.1 Proof of Theorem 8

In this section we give a proof of Theorem 8. We first prove the following observation about the Minkowski addition of down-monotone polytopes.

Observation 1. *Let $\mathcal{P} \subseteq [0,1]^d$ be any down-monotone polytope. Then for every $0 \leq a \leq b$, $a \cdot \mathcal{P} \subseteq b \cdot \mathcal{P}$. Let $\mathcal{P}_1, \mathcal{P}_2 \subseteq [0,1]^d$ both be down-monotone polytopes. Then for every $0 \leq a_1' \leq a_1$ and $0 \leq b_1' \leq b_1$, $a_1'\mathcal{P}_1 + b_1'\mathcal{P}_2 \subseteq a_1\mathcal{P}_1 + b_1\mathcal{P}_2$.*

Proof. For the first half of the statement, for every $\boldsymbol{x} \in a \cdot \mathcal{P}$, $\frac{\boldsymbol{x}}{a} \in \mathcal{P}$. Since $\mathcal{P}$ is down-monotone, $\frac{\boldsymbol{x}}{b} \in \mathcal{P}$. Thus $\boldsymbol{x} \in b \cdot \mathcal{P}$. As $a_1' \cdot \mathcal{P} \subseteq a_1 \cdot \mathcal{P}$ and $b_1' \cdot \mathcal{P} \subseteq b_1 \cdot \mathcal{P}$, the second half of the statement follows from the definition of the Minkowski addition. □

We use the following result from an unpublished manuscript by Cai et al. [CWD21]. A special case of the result appeared as Theorem 4 in [CDW12c] (conference version by the same authors). In particular, the result we use here is stated for a mixture of polytopes, while Theorem 4 in [CDW12c] is only for the polytope of all feasible reduced forms, but the proof is essentially the same. Interested readers are welcome to email the first author for a proof of Theorem 9.

Theorem 9 ([CWD21])**.** *Let ℓ be a positive integer. Let $\mathcal{P}$ be a mixture of $\{\mathcal{P}_i\}_{i\in[\ell]}$ over distribution $\mathcal{D} = \{q_i\}_{i\in[\ell]}$, where $\mathcal{P}_i \subseteq \mathbb{R}^d$ is a convex polytope for every $i \in [\ell]$. Assume for all i, the bit complexity of each corner of $\mathcal{P}_i$ is at most b. For any $\varepsilon > 0$ and integer $k \geq \Omega\left(d^4\left(b+\log\left(\frac{1}{\varepsilon}\right)\right)\right)$, let $\mathcal{D}'$ be the empirical distribution induced by $\lceil\frac{8kd}{\varepsilon^2}\rceil$ samples from $\mathcal{D}$. Let $\mathcal{P}$ be the mixture of $\{\mathcal{P}_i\}_{i\in[\ell]}$ over distribution $\mathcal{D}'$. With probability at least $1-2de^{-2dk}$ we have that*

1. *For all $\boldsymbol{x} \in \mathcal{P}$, there exists an $\boldsymbol{x}' \in \mathcal{P}'$ such that $||\boldsymbol{x}-\boldsymbol{x}'||_\infty \leq \varepsilon$.*
2. *For all $\boldsymbol{x}' \in \mathcal{P}'$, there exists an $\boldsymbol{x} \in \mathcal{P}$ such that $||\boldsymbol{x}-\boldsymbol{x}'||_\infty \leq \varepsilon$.*

To prove Theorem 8, we need the following lemmas.

Lemma 14. *Let $\mathcal{P} \subseteq [0,1]^d$ be a convex and down-monotone polytope. If $c\cdot\mathcal{P}^{box(\varepsilon)} \subseteq \mathcal{P}$ for some $\varepsilon > 0$ and $c \in (0,1]$, then $\frac{c}{2}\mathcal{P} \subseteq \frac{c}{2}\mathcal{P}^{tr(\varepsilon)} + \frac{c}{2}\mathcal{P}^{box(\varepsilon)} \subseteq \mathcal{P}$.*

Proof. First we prove that $\frac{c}{2}\mathcal{P} \subseteq \frac{c}{2}\mathcal{P}^{tr(\varepsilon)} + \frac{c}{2}\mathcal{P}^{box(\varepsilon)}$. Note that it is enough to prove that $\mathcal{P} \subseteq \mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)}$. For any $\boldsymbol{x} \in \mathcal{P}$, we consider the vectors $\boldsymbol{x}', \boldsymbol{x}'' \in [0,1]^d$ such that

$$\boldsymbol{x}_j' = \boldsymbol{x}_j \cdot \mathbb{1}[l_j(\mathcal{P}) \geq \varepsilon], \forall j \in [d]$$
$$\boldsymbol{x}_j'' = \boldsymbol{x}_j \cdot \mathbb{1}[l_j(\mathcal{P}) < \varepsilon], \forall j \in [d]$$

Note that $\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{x}''$. By the definition of $\mathcal{P}^{tr(\varepsilon)}$, $\boldsymbol{x}' \in \mathcal{P}^{tr(\varepsilon)}$. For $\boldsymbol{x}''$, we notice that for every $j \in [d]$, $\boldsymbol{x}_j'' = \boldsymbol{x}_j \cdot \mathbb{1}[l_j(\mathcal{P}) < \varepsilon] \leq l_j(\mathcal{P}) \cdot \mathbb{1}[l_j(\mathcal{P}) < \varepsilon]$. By the definition of $\mathcal{P}^{box(\varepsilon)}$, $\boldsymbol{x}'' \in \mathcal{P}^{box(\varepsilon)}$. Thus $\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{x}'' \in \mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)}$.

For the other direction, note that $\mathcal{P}^{tr(\varepsilon)} \subseteq \mathcal{P}$ by Definition 23 and $c\cdot\mathcal{P}^{box(\varepsilon)} \subseteq \mathcal{P}$ by assumption, so $\frac{c}{2}\mathcal{P}^{box(\varepsilon)} + \frac{c}{2}\mathcal{P}^{tr(\varepsilon)} \subseteq \frac{1}{2}\mathcal{P} + \frac{1}{2}\mathcal{P} = \mathcal{P}$. □

Lemma 15. *Let ℓ be a positive integer and $\mathcal{P} \subseteq [0,1]^d$ be a mixture of $\{\mathcal{P}_i\}_{i\in[\ell]}$ over distribution $\mathcal{D} = \{q_i\}_{i\in[\ell]}$, where for each $i \in [\ell]$, $\mathcal{P}_i \subseteq [0,1]^d$ is a convex and down-monotone polytope. Then $\mathcal{P}^{tr(\varepsilon)}$ is a mixture of $\{\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}$ over $\mathcal{D}$, where for each $i \in [\ell]$, $\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})} \subseteq \mathcal{P}_i$ is the $(\varepsilon,\mathcal{P})$-truncated polytope of $\mathcal{P}_i$ (Definition 23).*

Proof. To prove our statement, we first show that for each $\widehat{\boldsymbol{x}} \in \mathcal{P}^{tr(\varepsilon)}$, there exist $\{\widehat{\boldsymbol{x}}^{(i)} \in \mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}_{i\in[\ell]}$ such that $\widehat{\boldsymbol{x}} = \sum_{i\in[\ell]} q_i \widehat{\boldsymbol{x}}^{(i)}$. By definition of $\mathcal{P}^{tr(\varepsilon)}$, there exists $\boldsymbol{x} \in \mathcal{P}$ such that for each $j \in [d]$, $\widehat{\boldsymbol{x}}_j = \boldsymbol{x}_j \cdot \mathbb{1}[l_j(\mathcal{P}) \geq \varepsilon]$.

Since $\boldsymbol{x} \in \mathcal{P}$ and $\mathcal{P}$ is a mixture of $\{\mathcal{P}_i\}_{i\in[\ell]}$ over $\mathcal{D}$, there exist $\{\boldsymbol{x}^{(i)} \in \mathcal{P}_i\}_{i\in[\ell]}$ such that $\boldsymbol{x} = \sum_{i\in[\ell]} q_i \boldsymbol{x}^{(i)}$. For each $i \in [\ell]$, consider $\widehat{\boldsymbol{x}}^{(i)}$ be defined such that for all $j \in [d]$:

$$\widehat{\boldsymbol{x}}_j^{(i)} = \boldsymbol{x}_j^{(i)} \mathbb{1}[l_j(\mathcal{P}) \geq \varepsilon]$$

Clearly, $\widehat{\boldsymbol{x}}^{(i)} \in \mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}$ and $\widehat{\boldsymbol{x}} = \sum_{i\in[\ell]} q_i \widehat{\boldsymbol{x}}^{(i)}$. Similarly, we can argue that any point $\widehat{\boldsymbol{x}}$ that lies in the mixture of $\{\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}$ over $\mathcal{D}$ must also lie in $\mathcal{P}^{tr(\varepsilon)}$, which concludes the proof. □

Lemma 16. *Let ℓ be a positive integer, and $\mathcal{P} \subseteq [0,1]^d$ be a mixture of $\{\mathcal{P}_i\}_{i\in[\ell]}$ over distribution $\mathcal{D} = \{q_i\}_{i\in[\ell]}$, where $\mathcal{P}_i$ is a convex and down-monotone polytope in $[0,1]^d$ for every i. Then $\mathcal{P}$ is a convex and down-monotone polytope.*

Proof. For every $\boldsymbol{x} \in \mathcal{P}$, there exists a set of vectors $\{\boldsymbol{x}^{(i)}\}_{i\in[\ell]}$ such that $\boldsymbol{x}^{(i)} \in \mathcal{P}_i, \forall i$, and $\boldsymbol{x} = \sum_{i\in[\ell]} q_i \cdot \boldsymbol{x}^{(i)}$. Now consider any $\hat{\boldsymbol{x}}$ such that $\mathbf{0} \leq \hat{\boldsymbol{x}} \leq \boldsymbol{x}$. For each $i \in [\ell]$, let $\hat{\boldsymbol{x}}^{(i)} \in [0,1]^d$ be the vector such that $\hat{\boldsymbol{x}}_j^{(i)} = \boldsymbol{x}_j^{(i)} \cdot \hat{\boldsymbol{x}}_j/\boldsymbol{x}_j, \forall j \in [d]$. Clearly, $0 \leq \hat{x}_j^{(i)} \leq x_j^{(i)}$ for all $j \in [d]$. Since $\mathcal{P}_i$ is down-monotone, we have $\hat{\boldsymbol{x}}^{(i)} \in \mathcal{P}_i$. Note that for every $j \in [d]$, $\sum_i q_i \cdot \hat{\boldsymbol{x}}_j^{(i)} = (\sum_{i\in[\ell]} q_i \boldsymbol{x}_j^{(i)}) \cdot \hat{\boldsymbol{x}}_j/\boldsymbol{x}_j = \hat{\boldsymbol{x}}_j$. Thus $\hat{\boldsymbol{x}} = \sum_{i\in[\ell]} q_i \cdot \hat{\boldsymbol{x}}^{(i)} \in \mathcal{P}$. $\square$

To prove Theorem 8, we will also need the celebrated result of the equivalence between optimization and separation.

Theorem 10 ([KP80, GLS81])**.** *Let $\mathcal{P} \subseteq \mathbb{R}^d$ be a convex polytope and suppose we have access to an algorithm $\mathcal{A}(\boldsymbol{a}) : \mathbb{R}^d \to \mathcal{P}$, that takes input vector $\boldsymbol{a} \in \mathbb{R}^d$, outputs a vector $\boldsymbol{x}^* \in \mathcal{P}$ with bit complexity at most b, such that $\boldsymbol{x}^* \in argmax\{\boldsymbol{a} \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}\}$. Then we can construct a separation oracle SO for $\mathcal{P}$, where on any input $\boldsymbol{a} \in \mathbb{R}^d$ with bit complexity at most b', SO makes at most poly(d, b, b') queries to $\mathcal{A}$, and the input of each query has bit complexity no more than poly(d, b, b'). Moreover, the running time of SO on $\boldsymbol{a}$ is at most poly$(d, b, b', RT_{\mathcal{A}}(\text{poly}(d, b, b')))$. Here $RT_{\mathcal{A}}(c)$ is the running time of $\mathcal{A}$ with input whose bit complexity is at most c.*

Proof of Theorem 8:

Consider the polytopes $\mathcal{P}^{tr(\varepsilon)}$ and $\mathcal{P}^{box(\varepsilon)}$. By Lemma 15, $\mathcal{P}^{tr(\varepsilon)}$ is a mixture of $\{\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}_{i\in[\ell]}$ over distribution $\mathcal{D}$. Let $\widehat{\mathcal{D}}$ be the empirical distribution induced by $N = \lceil \frac{8kd}{\varepsilon^2} \rceil$ samples from $\mathcal{D}$. Let $\widehat{\mathcal{P}}^{tr(\varepsilon)}$ be the mixture of $\{\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}_{i\in[\ell]}$ over $\widehat{\mathcal{D}}$. By Theorem 9, we have that with probability at least $1 - 2de^{-2dk}$, both of the two following conditions hold:

1. For each $\widehat{p} \in \widehat{\mathcal{P}}^{tr(\varepsilon)}$, there exists a $p \in \mathcal{P}^{tr(\varepsilon)}$ such that $||p - \widehat{p}||_\infty \leq \varepsilon$.
2. For each $p \in \mathcal{P}^{tr(\varepsilon)}$, there exists a $\widehat{p} \in \widehat{\mathcal{P}}^{tr(\varepsilon)}$ such that $||p - \widehat{p}||_\infty \leq \varepsilon$.

For the rest of the proof, we condition on the event that both conditions hold. We consider the polytope $\widehat{\mathcal{P}} = \frac{c}{3}\left(\widehat{\mathcal{P}}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)}\right)$. First we are going to prove that $\widehat{\mathcal{P}} \subseteq \mathcal{P}$. By condition 1, we have that for any $\widehat{p}^{tr} \in \widehat{\mathcal{P}}^{tr(\varepsilon)}$, there exists a $p^{tr} \in \mathcal{P}^{tr(\varepsilon)}$ such that $||p^{tr} - \widehat{p}^{tr}||_\infty \leq \varepsilon$. Consider the vector $\tilde{p}^{tr}$ defined such that for each $j \in [d]$,

$$\tilde{p}_j^{tr} = \min\left(\widehat{p}_j^{tr}, p_j^{tr}\right).$$

Since for each $j \in [d]$, $\tilde{p}_j^{tr} \leq p_j^{tr}$ and $\mathcal{P}^{tr(\varepsilon)}$ is down-monotone (Lemma 16), we have $\tilde{p}^{tr} \in \mathcal{P}^{tr(\varepsilon)}$. Let vector $\tilde{p}^{box}$ be such that for every $j \in [d]$,

$$\tilde{p}_j^{box} = \widehat{p}_j^{tr} - \tilde{p}_j^{tr} = \widehat{p}_j^{tr} - \min(\widehat{p}_j^{tr}, p_j^{tr}).$$

Notice that for every $\boldsymbol{x} \in \mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}$, $x_j = 0$ for all j such that $l_j(\mathcal{P}) < \varepsilon$. Since $\mathcal{P}^{tr(\varepsilon)}$ and $\widehat{\mathcal{P}}^{tr(\varepsilon)}$ are both mixtures of $\{\mathcal{P}_i^{tr(\varepsilon,\mathcal{P})}\}_{i\in[\ell]}$, we have for every $\boldsymbol{x} \in \mathcal{P}^{tr(\varepsilon)}$ and $\widehat{\boldsymbol{x}} \in \widehat{\mathcal{P}}^{tr(\varepsilon)}$, $x_j = \widehat{x}_j = 0$ for all j such that $l_j(\mathcal{P}) < \varepsilon$. Therefore, we have $\tilde{p}_j^{box} \leq \varepsilon \cdot \mathbb{1}[l_j(\mathcal{P}) \geq \varepsilon] \leq \min(\varepsilon, l_j(\mathcal{P}))$, for every $j \in [d]$. The first inequality follows from the fact that $||p^{tr} - \widehat{p}^{tr}||_\infty \leq \varepsilon$, and that if $l_j(\mathcal{P}) < \varepsilon$, then $\widehat{p}_j^{tr} = p_j^{tr} = 0$. Thus $\tilde{p}^{box} \in \mathcal{P}^{box(\varepsilon)}$.

For every $\widehat{p}^{tr} \in \widehat{\mathcal{P}}^{tr(\varepsilon)}$, we have found $\tilde{p}^{tr} \in \mathcal{P}^{tr(\varepsilon)}$ and $\tilde{p}^{box} \in \mathcal{P}^{box(\varepsilon)}$ such that $\widehat{p}^{tr} = \tilde{p}^{tr} + \tilde{p}^{box}$. Thus

$$\begin{aligned}
&\widehat{\mathcal{P}}^{tr(\varepsilon)} \subseteq \mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)} \\
\Rightarrow &\widehat{\mathcal{P}}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)} \subseteq \mathcal{P}^{tr(\varepsilon)} + 2\mathcal{P}^{box(\varepsilon)} \subseteq \frac{3}{c} \cdot \mathcal{P} \\
\Rightarrow &\widehat{\mathcal{P}} = \frac{c}{3}\left(\widehat{\mathcal{P}}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)}\right) \subseteq \mathcal{P}
\end{aligned} \tag{10}$$

The second line follows from the assumption that $c\mathcal{P}^{box(\varepsilon)} \subseteq \mathcal{P}$ and $\mathcal{P}^{tr(\varepsilon)} \subseteq \mathcal{P}$ (by Definition 23 and the fact that $\mathcal{P}$ is down-monotone), and $c \leq 1$. Similarly, by switching the role of $\mathcal{P}^{tr(\varepsilon)}$ and $\widehat{\mathcal{P}}^{tr(\varepsilon)}$, with condition 2, we also have $\mathcal{P}^{tr(\varepsilon)} \subseteq \widehat{\mathcal{P}}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)}$. Thus

$$\begin{aligned}
&\mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)} \subseteq \widehat{\mathcal{P}}^{tr(\varepsilon)} + 2\mathcal{P}^{box(\varepsilon)} \\
\Rightarrow &\mathcal{P} \subseteq \mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)} \subseteq 2\left(\widehat{\mathcal{P}}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)}\right) = \frac{6}{c}\widehat{\mathcal{P}} \\
\Rightarrow &\frac{c}{6}\mathcal{P} \subseteq \widehat{\mathcal{P}}
\end{aligned}$$

The second line follows from $\mathcal{P} \subseteq \mathcal{P}^{tr(\varepsilon)} + \mathcal{P}^{box(\varepsilon)}$ (Lemma 14), the origin $\mathbf{0} \in \widehat{\mathcal{P}}^{tr(\varepsilon)}$, and the definition of $\widehat{\mathcal{P}}$. Thus $\frac{c}{6}\mathcal{P} \subseteq \widehat{\mathcal{P}} \subseteq \mathcal{P}$.

To construct a separation oracle for $\widehat{\mathcal{P}}$, it is sufficient to optimize any linear objective over $\widehat{\mathcal{P}}$. For every $\boldsymbol{a} \in \mathbb{R}^d$, we are going to solve the maximization problem $\max\{\boldsymbol{a} \cdot \boldsymbol{x} : \boldsymbol{x} \in \widehat{\mathcal{P}}\}$.

Let $\{i_1, ..., i_N\}$ be the N samples from $\mathcal{D}$, where $i_k \in [\ell]$ for $k \in [N]$. We notice that

$$\widehat{\mathcal{P}} = \sum_{k\in[N]} \frac{c}{3N} \cdot \mathcal{P}_{i_k}^{tr(\varepsilon,\mathcal{P})} + \frac{c}{3}\mathcal{P}^{box(\varepsilon)}$$

is the Minkowski addition of a set of polytopes. Thus in order to maximize over $\widehat{\mathcal{P}}$, it's sufficient to maximize over each polytope. In other words, it is sufficient to solve $\max\{\boldsymbol{a} \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}^{box(\varepsilon)}\}$ and $\max\{\boldsymbol{a} \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_{i_k}^{tr(\varepsilon,\mathcal{P})}\}$ for each $k \in [N]$. First consider $\mathcal{P}^{box(\varepsilon)}$. Since the polytope is a "box" where the constraint for each coordinate j is separate: $x_j \leq \min(\varepsilon, l_j(\mathcal{P}))$. Thus the optimum $\boldsymbol{x}^{box} \in \mathcal{P}^{box(\varepsilon)}$ satisfies that $x_j^{box} = \min(l_j(\mathcal{P}), \varepsilon) \cdot \mathbb{1}[a_j > 0]$. Thus the optimum $\boldsymbol{x}^{box}$ can be computed in time $O(d \cdot (b + \log 1/\varepsilon + b'))$ and its bit complexity is at most $O(d \cdot (b + \log 1/\varepsilon))$, where b' is the bit complexity of $\boldsymbol{a}$.

Now we show how to solve $\max\{\boldsymbol{a} \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_{i_k}^{tr(\varepsilon,\mathcal{P})}\}$ using a single query to $\mathcal{Q}_{i_k}(\cdot)$, for every $k \in [N]$. Consider the vector $\boldsymbol{a}' \in \mathbb{R}^d$ such that $a'_j = a_j \cdot \mathbb{1}[l_j(\mathcal{P}) \geq \varepsilon], \forall j \in [d]$. Then clearly

$$\max\{\boldsymbol{a} \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_{i_k}^{tr(\varepsilon,\mathcal{P})}\} = \max\{\boldsymbol{a}' \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_{i_k}\}$$

Let $\widehat{\boldsymbol{x}}^{(i_k)}$ be the output from oracle $\mathcal{Q}_{i_k}(\boldsymbol{a}')$, then $\widehat{\boldsymbol{x}}^{(i_k)} \in \text{argmax}\{\boldsymbol{a}' \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_{i_k}\}$. Consider the element $\boldsymbol{x}^{(i_k)} \in [0,1]^d$ such that for each $j \in [d]$, $x_j^{(i_k)} = \widehat{x}_j^{(i_k)} \cdot \mathbb{1}[l_j(\mathcal{P}) \geq \varepsilon]$. Then $\boldsymbol{x}^{(i_k)} \in \text{argmax}\{\boldsymbol{a}' \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathcal{P}_{i_k}^{tr(\varepsilon,\mathcal{P})}\}$ and its bit complexity is at most the bit complexity of $\widehat{\boldsymbol{x}}^{(i_k)}$, which is at most b, by our assumption on $\mathcal{Q}_i(\cdot)$.

Thus we have

$$\sum_{k\in[N]} \frac{c}{3N} \cdot \boldsymbol{x}^{(i_k)} + \frac{c}{3} \cdot \boldsymbol{x}^{box} \in \text{argmax}\{\boldsymbol{a} \cdot \boldsymbol{x} : \boldsymbol{x} \in \widehat{\mathcal{P}}\}$$

To sum up, we provide an algorithm to optimize any linear objective over $\widehat{\mathcal{P}}$. Moreover, the output of our optimization algorithm always has bit complexity $\text{poly}(b, k, d, 1/\varepsilon)$. For any $\boldsymbol{a} \in \mathbb{R}^d$ with bit complexity b', our optimization algorithm runs in time $\text{poly}(b, b', k, d, 1/\varepsilon)$ and make $N = \left\lceil \frac{8kd}{\varepsilon^2} \right\rceil$ queries to the oracles in $\{\mathcal{Q}_i\}_{i\in[\ell]}$. Using Theorem 10, we can construct a separation oracle for $\widehat{\mathcal{P}}$ that satisfies the properties in the statement of Theorem 8.

□

D.2 Approximating Single-Bidder Marginal Reduced Form Polytope for Constrained-Additive Buyers

In this section, we prove Theorem 2 using Theorem 8. The goal is to show that the single-bidder marginal reduced form polytope W_i satisfies the requirements of Theorem 8. Recall that for every buyer i, the support of her type is $\mathcal{T}_i$ and the support of her value for each item j is $\mathcal{T}_{ij}$. Additionally, W_i is a subset of $[0,1]^{\sum_{j\in[m]}|\mathcal{T}_{ij}|}$, and there is a coordinate for every $j \in [m]$ and every $t_{ij} \in \mathcal{T}_{ij}$. To ease the notation, we will use t_{ij}'s to index the coordinates throughout this section. Since each buyer is constrained-additive, we denote $\mathcal{F}_i$ the feasibility constraint of buyer i, and drop the subscript if the buyer is fixed or clear from context.

Lemma 17. *For each $i \in [n], j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$, $l_{t_{ij}}(W_i) = f_{ij}(t_{ij})$. Recall that $l_{t_{ij}}(W_i)$ is the width of W_i at coordinate t_{ij} (Definition 23).*

Proof. For every $i \in [n]$ and every $\widehat{w}_i \in W_i$, by Definition 5, there exists a number $\sigma_S(t_i) \in [0,1]$ for every $t_i \in \mathcal{T}_i, S \in \mathcal{F}_i$ such that

1. $\sum_{S\in\mathcal{F}_i} \sigma_S(t_i) \le 1, \forall t_i \in \mathcal{T}_i$.
2. $\widehat{w}_{ij}(t_{ij}) = f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S\in\mathcal{F}_i: j\in S} \sigma_S(t_i)$, for all $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$.

Thus for every j, t_{ij}, by combining both properties above, we have

$$\widehat{w}_{ij}(t_{ij}) \le f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot 1 = f_{ij}(t_{ij}).$$

Moreover, for every j, t_{ij}, choosing $\widehat{\sigma}$ such that: [25]

$$\widehat{\sigma}_S(t_i') = \begin{cases} 1 & \text{if } t_{ij}' = t_{ij} \wedge S = \{j\} \\ 0 & \text{o.w.} \end{cases}$$

induces an element $\widetilde{w}_i \in W_i$ such that $\widetilde{w}_{ij}(t_{ij}) = f_{ij}(t_{ij})$. Thus $l_{t_{ij}}(W_i) = f_{ij}(t_{ij})$. □

Definition 25. *For any buyer i and type $t_i \in \mathcal{T}_i$, consider $W_{t_i} \subseteq [0,1]^{\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ defined as follows: $w_i \in W_{t_i}$ if and only if there exists a collection of non-negative numbers $\{\sigma_S\}_{S\in\mathcal{F}_i}$ such that*

1. $\sum_{S\in\mathcal{F}_i} \sigma_S \le 1$.
2. $w_{ij}(t_{ij}') = \sum_{S\in\mathcal{F}_i, j\in S} \sigma_S \cdot \mathbb{1}[t_{ij}' = t_{ij}], \forall j \in [m], t_{ij}' \in \mathcal{T}_{ij}$.

The following observation directly follows from Definition 5 and Definition 25.

Observation 2. *W_i is mixture of $\{W_{t_i}\}_{t_i\in\mathcal{T}_i}$ over distribution D_i. Recall that D_i is the distribution for buyer i's type t_i.*

Lemma 18. *For every i and $t_i \in \mathcal{T}_i$, W_{t_i} is a convex and down-monotone polytope. Moreover, given access to a demand oracle $\text{DEM}_i(t_i, \cdot)$ for buyer i, we can calculate an element*

$$w_i^* \in \mathit{argmax}_{w_i\in W_{t_i}} \boldsymbol{a} \cdot w_i,$$

for any $\boldsymbol{a} \in \mathbb{R}^{\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ with a single query to the demand oracle. Moreover, the bit complexity of w_i^* is at most $\sum_{j\in[m]}|\mathcal{T}_{ij}|$.

[25] $\widehat{\sigma}$ is simply the allocation that gives buyer i item j when her value for item j is t_{ij} and does not give her anything otherwise.

Proof. For every feasible set $S \in \mathcal{F}_i$, consider allocation $\lambda^{(S)} = \{\mathbb{1}[S' = S]\}_{S' \in \mathcal{F}_i}$. The set of $\sigma = \{\sigma_S\}_{S \in \mathcal{F}_i}$ that satisfy property 1 of Definition 25 is equivalent to the set of all convex combinations of the $\lambda^{(S)}$'s and the origin $\mathbf{0}$. More specifically, $\sigma = \sum_{S \in \mathcal{F}_i} \sigma_S \cdot \lambda^{(S)} + (1 - \sum_{S \in \mathcal{F}_i} \sigma_S) \cdot \mathbf{0}$. Hence, the set of σ's is a convex polytope P in $[0,1]^{|\mathcal{F}_i|}$. Since W_{t_i} is a projection of P to $[0,1]^{\sum_{j\in[m]} |\mathcal{T}_{ij}|}$, W_{t_i} is also a convex polytope.

Next, we prove W_{t_i} is down-monotone. Consider any $w_i \in W_{t_i}$. By Definition 25, $w_{ij}(t'_{ij}) = 0$ for all j, t'_{ij} such that $t'_{ij} \neq t_{ij}$. To prove that W_{t_i} is down-monotone, it suffices to prove that for every $j_0 \in [m]$, any vector $\widetilde{w}_i$, achieved by only decreasing the coordinate t_{ij_0} from w_i, is still in W_{t_i}. Formally, $\widetilde{w}_i$ satisfies

- $\widetilde{w}_{ij_0}(t_{ij_0}) < w_{ij_0}(t_{ij_0})$.
- $\widetilde{w}_{ij}(t_{ij}) = w_{ij}(t_{ij})$, if $j \neq j_0$.
- $\widetilde{w}_{ij}(t'_{ij}) = 0$, for any $j \in [m]$, if $t'_{ij} \neq t_{ij}$.

Let $\{\sigma_S\}_{S \in \mathcal{F}_i}$ be the vector of numbers associated with w_i in Definition 25. Let $\alpha = \frac{\widetilde{w}_{ij_0}(t_{ij_0})}{w_{ij_0}(t_{ij_0})} < 1$. Consider another vector $\{\widetilde{\sigma}_S\}_{S \in \mathcal{F}_i}$ where for every $S \in \mathcal{F}_i$,

$$\widetilde{\sigma}_S = \begin{cases} \alpha \cdot \sigma_S + (1-\alpha) \cdot (\sigma_{S \cup \{j_0\}} + \sigma_S), & j_0 \notin S \wedge S \cup \{j_0\} \in \mathcal{F}_i \\ \alpha \cdot \sigma_S, & j_0 \in S \\ \sigma_S, & \text{o.w.} \end{cases}$$

We notice that the above definition is well-defined because $S \in \mathcal{F}_i$ as long as $S \cup \{j_0\} \in \mathcal{F}_i$. Also $\sum_{S \in \mathcal{F}_i} \widetilde{\sigma}_S \leq 1$. Intuitively, $\sigma = \{\sigma_S\}_{S \in \mathcal{F}_i}$ represents a randomized allocation of sets $S \in \mathcal{F}_i$ to the bidder. Then $\{\widetilde{\sigma}_S\}_{S \in \mathcal{F}_i}$ represents another randomized allocation: choose a set S according to the the randomized allocation σ, if S contains j_0, then throw away j_0 with probability $1-\alpha$.

Now we have $\sum_{S \in \mathcal{F}_i, j_0 \in S} \widetilde{\sigma}_S = \alpha \sum_{S \in \mathcal{F}_i, j_0 \in S} \sigma_S = \widetilde{w}_{ij_0}(t_{ij_0})$ and $\sum_{S \in \mathcal{F}_i, j \in S} \widetilde{\sigma}_S = w_{ij}(t_{ij})$, for all $j \neq j_0$. Thus $\widetilde{w}_i \in W_{t_i}$, and W_{t_i} is down-monotone.

It remains to show that we can find an element in $\text{argmax}_{w_i \in W_{t_i}} \boldsymbol{a} \cdot w_i$ for any $\boldsymbol{a} \in \mathbb{R}^{\sum_{j\in[m]} |\mathcal{T}_{ij}|}$, given access to a demand oracle. W.l.o.g. we assume that for each $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$, $a_{ij}(t_{ij}) \leq \frac{1}{2} \cdot t_{ij}$ (by scaling and the fact that $t_{ij} > 0$). Observe that any corner w_i of the polytope W_{t_i} corresponds to the choice of $\{\sigma_S\}_{S \in \mathcal{F}_i}$ such that $\sigma_S = \mathbb{1}[S = T]$ for some particular $T \in \mathcal{F}_i$, i.e. $w_{ij}(t'_{ij}) = \mathbb{1}[j \in T \wedge t'_{ij} = t_{ij}]$.

Since $w_i^* \in \text{argmax}_{w_i \in W_{t_i}} \boldsymbol{a} \cdot w_i$ is a corner of W_{t_i}. We have

$$\begin{aligned} &\max\nolimits_{w_i \in W_{t_i}} \boldsymbol{a} \cdot w_i \\ =&\max_{S \in \mathcal{F}_i} \sum_{j \in S} a_{ij}(t_{ij}) \\ =&\max_{S \in \mathcal{F}_i} \sum_{j \in S} a_{ij}(t_{ij})^+ \\ =&\max_{S \in \mathcal{F}_i} \sum_{j \in S} (t_{ij} - (t_{ij} - a_{ij}(t_{ij})^+)) \end{aligned}$$

Here $x^+ = \max\{x, 0\}$. The second equality holds because $\mathcal{F}_i$ is downward-closed. Notice that $a_{ij}(t_{ij})^+ \leq \frac{1}{2} t_{ij}$. Thus with a single query to the demand oracle, with type t_i and prices $p_{ij} = t_{ij} - a_{ij}(t_{ij})^+ \geq 0, \forall j$, we can find $\text{argmax}_{w_i \in W_{t_i}} \boldsymbol{a} \cdot w_i$. The bit complexity of w_i^* is at most $\sum_{j \in [m]} |\mathcal{T}_{ij}|$. $\square$

Lemma 19. *Let $T = \sum_{j\in[m]} |\mathcal{T}_{ij}|$. For any $\varepsilon < \frac{1}{T}$, $(1 - \varepsilon T) W_i^{box(\varepsilon)} \subseteq W_i$.*

Proof. For any $w_i \in (1-\varepsilon T)W_i^{box(\varepsilon)}$, we prove that $w_i \in W_i$.

Consider the following set of numbers $\{\sigma_S(t_i)\}_{t_i,S}$ (see Definition 5): For each $j \in [m], t_{ij} \in \mathcal{T}_{ij}$, let $c_j(t_{ij}) = \min\left(\frac{\varepsilon}{f_{ij}(t_{ij})}, 1\right)$ and

$$p_j(t_{ij}) = \frac{w_{ij}(t_{ij})}{f_{ij}(t_{ij})c_j(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \prod_{j' \neq j}(1 - c_{j'}(t_{ij'}))}.$$

Note that for every $j' \in [m]$, there exists a value $t_{ij'} \in \mathcal{T}_{ij'}$ such that $f_{ij'}(t_{ij'}) \geq 1/|\mathcal{T}_{ij'}|$. Due to our choice of ε, the corresponding $c_{j'}(t_{ij'}) < 1$. Hence, $\sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \prod_{j' \neq j}(1 - c_{j'}(t_{ij'})) > 0$, and $p_j(t_{ij})$ is well-defined.

For every t_i, define

$$\sigma_S(t_i) = \begin{cases} p_j(t_{ij}) \cdot c_j(t_{ij}) \cdot \prod_{j' \neq j}(1 - c_{j'}(t_{ij'})), & \text{if } S = \{j\} \text{ for some } j \in [m] \\ 0, & \text{o.w.} \end{cases}$$

For every j and t_{ij}, let $C_j(t_{ij})$ be the independent Bernoulli random variable that is 1 with probability $c_j(t_{ij})$. Then for every j,

$$\Pr_{C_j(t_{ij}), t_{ij} \sim D_{ij}}\left[C_j(t_{ij}) = 1\right] = \sum_{t_{ij} \in \mathcal{T}_{ij}} f_{ij}(t_{ij}) \cdot \min\left(\frac{\varepsilon}{f_{ij}(t_{ij})}, 1\right) \leq \varepsilon \cdot |\mathcal{T}_{ij}|$$

By the union bound,

$$\begin{aligned}
\sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \prod_{j' \neq j}(1 - c_{j'}(t_{ij'})) &= \Pr_{\substack{t_{i,-j} \sim D_{i,-j} \\ \forall j' \neq j, C_k(t_{ij'})}}\left[C_{j'}(t_{ij'}) = 0, \forall j' \neq j\right] \\
&= 1 - \Pr_{\substack{t_{i,-j} \sim D_{i,-j} \\ \forall j' \neq j, C_k(t_{ij'})}}\left[C_{j'}(t_{ij'}) = 1, \exists j' \neq j\right] \\
&\geq 1 - \sum_{j \neq j'} \Pr_{\substack{t_{ij} \sim D_{ij} \\ C_j(t_{ij})}}\left[C_j(t_{ij}) = 1\right] \\
&\geq 1 - \varepsilon \cdot T
\end{aligned}$$

Now we show that $w_i \in W_i$ by verifying both properties in Definition 5. For the first property, since $w_i \in (1-\varepsilon T)W_i^{box(\varepsilon)}$, $0 \leq w_{ij}(t_{ij}) \leq (1 - \varepsilon \cdot T) \cdot \min\{\varepsilon, l_{t_{ij}}(W_i)\} = (1 - \varepsilon \cdot T) \cdot f_{ij}(t_{ij}) \cdot c_j(t_{ij})$. The equality is due to the definition of $c_j(t_{ij})$ and Lemma 17. Thus $p_j(t_{ij}) \leq 1$ for every j and t_{ij}. We have that $\sum_S \sigma_S(t_i) = \sum_{j \in [m]} p_j(t_{ij}) \cdot c_j(t_{ij}) \cdot \prod_{j' \neq j}(1 - c_{j'}(t_{ij'})) \leq \sum_{j \in [m]} c_j(t_{ij}) \cdot \prod_{j' \neq j}(1 - c_{j'}(t_{ij'})) \leq \prod_{j \in [m]}\left(c_j(t_{ij}) + (1 - c_j(t_{ij}))\right) = 1$.

The second property:

$$\begin{aligned}
&f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S: j \in S} \sigma_S(t_{ij}, t_{i,-j}) \\
=&f_{ij}(t_{ij}) \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot p_j(t_{ij}) \cdot c_j(t_{ij}) \cdot \prod_{j' \neq j}(1 - c_{j'}(t_{ij'})) = w_{ij}(t_{ij})
\end{aligned}$$

Thus by Definition 5, $w_i \in W_i$.

□

Proof of Theorem 2:

We simply verify that W_i satisfies all assumptions in Theorem 8. Recall that $T = \sum_{i,j} |\mathcal{T}_{ij}|$ and b is an upper bound of the bit complexity of all $f_{ij}(t_{ij})$'s and all t_{ij}'s. We have the following:

1. By Observation 2, we have that W_i is a mixture of $\{W_{t_i}\}_{t_i \in \mathcal{T}_i}$ over distribution D_i.

2. By Lemma 18, for each $t_i \in \mathcal{T}_i$, W_{t_i} is a convex and down-monotone polytope. Given access to the demand oracle $\text{DEM}_i(t_i, \cdot)$, we can find an element in $\arg\max\{\boldsymbol{a} \cdot w_i : w_i \in W_{t_i}\}$ in time $poly(b', b, T)$ and a single query to $\text{DEM}_i(t_i, \cdot)$, where b' is the bit complexity of the input $\boldsymbol{a}$. Note that each output of the demand oracle has bit complexity at most T.

3. By Lemma 17, $l_{t_{ij}}(W_i) = f_{ij}(t_{ij})$,for all j and t_{ij}. Thus, each $l_{t_{ij}}(W_i)$ has bit complexity at most b.

4. By Lemma 19, $(1 - \varepsilon T) W_i^{box(\varepsilon)} \subseteq W_i$ for any $\varepsilon < \frac{1}{T}$. Choosing $\varepsilon = \frac{1}{2T}$ obtains $\frac{1}{2} W_i^{box(\frac{1}{2T})} \subseteq W_i$.

For any $\delta \in (0, 1)$, we apply Theorem 8 with parameter $k = poly(n, m, T, b, \log(1/\delta))$, $c = \frac{1}{2}$ and $\varepsilon = \frac{1}{2T}$. The probability that the algorithm successfully constructs a polytope that satisfies both properties of Theorem 8 is at least $1 - \delta$. We have $\frac{1}{12} \cdot W_i \subseteq \widehat{W}_i \subseteq W_i$ by the first property of Theorem 8 with $c = \frac{1}{2}$. Since the vertex-complexity of W_{t_i} for each t_i is no more than T, and the vertex-complexity for $W_i^{box(\frac{1}{2T})}$ is no more than $\text{poly}(b, T)$, the vertex-complexity for $\widehat{W}_i$ is no more than $poly(n, m, T, b, \log(1/\delta))$. The running time of the algorithm and the separation oracle SO for $\widehat{W}_i$ follows from the second property of Theorem 8.
□

At last, we give the proof of Theorem 7 by combining Theorem 3 and Theorem 2.

Proof of Theorem 7: Fix any $\delta \in (0, 1)$. Recall that in the LP of Figure 3, we use an estimation of PREV, which is $\widetilde{\text{PREV}}$, according to Lemma 2. Denote $\mathcal{E}_1$ the event that an RPP mechanism is successfully computed and $\mathcal{E}_2$ the event that the algorithm in Theorem 2 successfully constructs a convex polytope $\widehat{W}_i$ that satisfies both properties in the statement of Theorem 2 for each buyer $i \in [n]$. Note that $\mathcal{E}_1$ happens with probability at least $1 - \frac{2}{nm}$ and we take enough samples to make sure that $\mathcal{E}_2$ happens with probability at least $1 - \delta$, by the union bound, the probability that both $\mathcal{E}_1$ and $\mathcal{E}_2$ happen is at least $1 - \delta - \frac{2}{nm}$. From now on, we will condition on both events $\mathcal{E}_1$ and $\mathcal{E}_2$.

Now in the LP of Figure 3, we replace W_i by $\widehat{W}_i$ for every i (we will call it the modified LP). By property 1 and 2 of Theorem 2, we can solve the modified LP using the separation oracle for $\widehat{W}_i$, in time $\text{poly}(b, n, m, T, \log(1/\delta))$ (recall that $T = \sum_{i,j} |\mathcal{T}_{ij}|$) according to Theorem 5. Let $x^* = (w^*, \lambda^*, \hat{\lambda}^*, \mathbf{d}^*)$ be an optimal solution of the modified LP. Then $w_i^* \in \widehat{W}_i$ for every i according to Constraint (1) in the modified LP. By property 1 of Theorem 2, we have $\widehat{W}_i \subseteq W_i$. Thus $w_i^* \in W_i, \forall i$ and hence x^* is also a feasible solution of the original LP in Figure 3.

Recall that OPT_{LP} is the optimum objective of the original LP. Denote OPT'_{LP} the optimum objective of the modified LP. Thus in order to prove that $(w^*, \lambda^*, \hat{\lambda}^*, \mathbf{d}^*)$ is an approximately-optimal solution in the original LP, it suffices to show that $\text{OPT}'_{\text{LP}} \geq c \cdot \text{OPT}_{\text{LP}}$. Take any feasible solution $(w, \lambda, \hat{\lambda}, \mathbf{d})$ of the original LP. We have that $w_i \in W_i$ for every i. Now consider another set of variables $(w', \lambda', \hat{\lambda}, \mathbf{d})$ such that $w'_{ij}(t_{ij}) = c \cdot w_{ij}(t_{ij})$ and $\lambda'_{ij}(t_{ij}, \beta_{ij}, \delta_{ij}) = c \cdot \lambda_{ij}(t_{ij}, \beta_{ij}, \delta_{ij})$, for all $i, j, t_{ij}, \beta_{ij}, \delta_{ij}$, where $c = 1/12$. We verify that $(w', \lambda', \hat{\lambda}, \mathbf{d})$ is a feasible solution for the modified LP. For Constraint (1), since $c \cdot W_i \subseteq \widehat{W}_i$ (property 1 of Theorem 2), we have that $w'_i \in \widehat{W}_i$, for all i. For Constraint (3), it holds since we multiply both λ and w by c. Constraint (2), (4) and (7) hold, as for each of them their LHS is smaller while the RHS remains unchanged. Every other constraint holds since both of their LHS and RHS remain the same. Thus $(w', \lambda', \hat{\lambda}, \mathbf{d})$ is a feasible solution for the modified LP.

Now notice that the objective of the solution $(w', \lambda', \hat{\lambda}, \mathbf{d})$ is exactly c times the objective of the solution $(w, \lambda, \hat{\lambda}, \mathbf{d})$. By choosing $(w, \lambda, \hat{\lambda}, \mathbf{d})$ to be the optimal solution of the original LP, we have that $\text{OPT}'_{\text{LP}} \geq c \cdot \text{OPT}_{\text{LP}}$. It implies that the objective of $(w^*, \lambda^*, \hat{\lambda}^*, \mathbf{d}^*)$ is at least $c \cdot \text{OPT}_{\text{LP}}$. Thus if we compute the simple mechanisms using the decision variables $(w^*, \lambda^*, \hat{\lambda}^*, \mathbf{d}^*)$. By Theorem 2, we have

$$c_1 \cdot \text{REV}(\mathcal{M}_{\text{PP}}) + c_2 \cdot \text{REV}(\mathcal{M}_{\text{TPT}}) \geq \text{OPT}'_{\text{LP}} \geq c \cdot \text{OPT}_{\text{LP}} \geq c \cdot \text{OPT}$$

We finish our proof by noticing that the simple mechanisms can be computed in time $\text{poly}(n, m, T)$ given the solution $(w^*, \lambda^*, \hat{\lambda}^*, \mathbf{d}^*)$. □

E Accessing Single-Bidder Marginal Reduced Form Polytopes for XOS Valuations

Our goal in this section is to prove Theorem 1 for XOS valuations.

Theorem 11. *(Restatement of Theorem 1 for XOS valuations) Let $T = \sum_{i,j} |\mathcal{T}_{ij}|$ and b be the bit complexity of the problem instance (Definition 3). For XOS buyers, for any $\delta > 0$, there exists an algorithm that computes a rationed posted price mechanism or a two-part tariff mechanism, such that the revenue of the mechanism is at least $c \cdot \text{OPT}$ for some absolute constant $c > 0$ with probability $1 - \delta - \frac{2}{nm}$. Our algorithm assumes query access to a value oracle and an adjustable demand oracle (see Section 2) of buyers' valuations, and has running time $\text{poly}(n, m, T, b, \log(1/\delta))$.*

We remind the readers the definition of the single-bidder marginal reduced form polytope W_i for XOS valuations:

Definition 26 (Restatement of Definition 14)**.** *For every $i \in [n]$, the single-bidder marginal reduced form polytope $W_i \subseteq [0,1]^{2 \cdot \sum_j |\mathcal{T}_{ij}|}$ is defined as follows. Let $\pi_i = (\pi_{ij}(t_{ij}))_{j, t_{ij} \in \mathcal{T}_{ij}}$ and $w_i = (w_{ij}(t_{ij}))_{j, t_{ij} \in \mathcal{T}_{ij}}$. Then $(\pi_i, w_i) \in W_i$ if and only if there exist a number $\sigma_S^{(k)}(t_i) \in [0,1]$ for every $t_i \in \mathcal{T}_i, S \subseteq [m], k \in [K]$, such that*

1. $\sum_{S,k} \sigma_S^{(k)}(t_i) \leq 1, \forall t_i \in \mathcal{T}_i$.
2. $\pi_{ij}(t_{ij}) = f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S: j \in S} \sum_{k \in [K]} \sigma_S^{(k)}(t_{ij}, t_{i,-j})$, *for all* $i, j, t_{ij} \in \mathcal{T}_{ij}$.
3. $w_{ij}(t_{ij}) \leq f_{ij}(t_{ij}) \cdot \sum_{t_{i,-j}} f_{i,-j}(t_{i,-j}) \cdot \sum_{S: j \in S} \sum_{k \in [K]} \sigma_S^{(k)}(t_{ij}, t_{i,-j}) \cdot \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})}$, *for all* $i, j, t_{ij} \in \mathcal{T}_{ij}$. *If $V_{ij}(t_{ij}) = 0$, we slightly abuse notation and treat $\frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})}$ as 0 for all k.*

The following observation follows directly from Definition 14.

Observation 3. *For any $(\pi_i, w_i) \in W_i$ and any $j \in [m], t_{ij} \in \mathcal{T}_{ij}$, $w_{ij}(t_{ij}) \leq \pi_{ij}(t_{ij}) \leq f_{ij}(t_{ij})$.*

Proof. The first inequality follows directly from the fact that $\alpha_{ij}^{(k)}(t_{ij}) \leq V_{ij}(t_{ij}) = \max_{k'} \alpha_{ij}^{(k')}(t_{ij})$. The second inequality follows from $\sum_{S: j \in S} \sum_{k \in [K]} \sigma_S^{(k)}(t_{ij}, t_{i,-j}) \leq \sum_S \sum_k \sigma_S^{(k)}(t_{ij}, t_{i,-j}) \leq 1$. □

In Appendix C.3, we provide an LP (Figure 4) that helps us to compute the simple mechanisms efficiently. In Theorem 6, we have proved that given any optimal solution to the LP in Figure 4, we can compute a simple mechanism in polynomial time, whose revenue is a constant factor of the optimal revenue. However, constraint (1) is implicit and thus it's unclear if we can solve the LP in polynomial time. Similar to the idea in Appendix D, we fix this issue by constructing another polytope $\widehat{W}_i$. Unfortunately, for XOS valuations, W_i is not a down-monotone polytope anymore. To see this, we simply notice that by Observation 3, $w_{ij}(t_{ij}) \leq \pi_{ij}(t_{ij})$ for every coordinate (j, t_{ij}). Thus given any $(\pi_i, w_i) \in W_i$ where $w_{ij}(t_{ij}) > 0$ for some j, t_{ij}, the vector $(\mathbf{0}, w_i)$ is clearly not in W_i. Thus the argument in Appendix D does not apply here.

E.1 Basic Properties of the Single-Bidder Marginal Reduced Form for XOS Valuations

In this section we present some basic definitions and properties of the single-bidder Marginal Reduced Form polytope W_i (Definition 14). We fixed any buyer i throughout this section unless otherwise specified.

Definition 27. *For any $\varepsilon > 0$, we denote as $W_i^{tr(\varepsilon)} \subseteq [0,1]^{2\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ the ε-truncated polytope of W_i. An element $(\widehat{\pi}_i, \widehat{w}_i) \in W_i^{tr(\varepsilon)}$ if there exists $(\pi_i, w_i) \in W_i$ such that for all $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$:*

$$\widehat{w}_{ij}(t_{ij}) = w_{ij}(t_{ij}) \cdot \mathbb{1}[f_{ij}(t_{ij}) \geq \varepsilon]$$
$$\widehat{\pi}_{ij}(t_{ij}) = \pi_{ij}(t_{ij}) \cdot \mathbb{1}[f_{ij}(t_{ij}) \geq \varepsilon]$$

Similar to Section D, we show that $W_i^{tr(\varepsilon)}$ is a mixture (Definition 22) of a set of polytopes $\{W_{t_i}^{tr(\varepsilon)}\}_{t_i\in\mathcal{T}_i}$ defined in Definition 28 over $\mathcal{D}$.

Definition 28. *For any $i, t_i \in \mathcal{T}_i$ and $\varepsilon > 0$, we define the polytopes $W_{t_i}, W_{t_i}^{tr(\varepsilon)} \subseteq [0,1]^{2\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ as follows: An element $\left(x = \left\{x(t'_{ij})\right\}_{t'_{ij}\in\mathcal{T}_{ij}}, y = \left\{y(t'_{ij})\right\}_{t'_{ij}\in\mathcal{T}_{ij}}\right) \in W_{t_i}$ if there exists a collection of non-negative numbers $\{\sigma_S^{(k)}\}_{S\subseteq[m],k\in[K]}$, such that $\sum_{S\subseteq[m]}\sum_{k\in[K]}\sigma_S^{(k)} \leq 1$, and for any $j, t'_{ij} \in \mathcal{T}_{ij}$,*

$$x(t'_{ij}) = \sum_{S:j\in S}\sum_{k\in[K]}\sigma_S^{(k)} \cdot \mathbb{1}[t'_{ij} = t_{ij}]$$

$$y(t'_{ij}) \leq \sum_{S:j\in S}\sum_{k\in[K]}\sigma_S^{(k)} \cdot \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})} \cdot \mathbb{1}[t'_{ij} = t_{ij}].$$

Moreover, an element $\left(\hat{x} = \left\{\hat{x}(t'_{ij})\right\}_{t'_{ij}\in\mathcal{T}_{ij}}, \hat{y} = \left\{\hat{y}(t'_{ij})\right\}_{t'_{ij}\in\mathcal{T}_{ij}}\right) \in W_{t_i}^{tr(\varepsilon)}$ if there exists $(x,y) \in W_{t_i}$ such that for any j, t'_{ij},

$$\hat{x}(t'_{ij}) = x(t'_{ij}) \cdot \mathbb{1}[f_{ij}(t_{ij}) \geq \varepsilon]$$
$$\hat{y}(t'_{ij}) = y(t'_{ij}) \cdot \mathbb{1}[f_{ij}(t_{ij}) \geq \varepsilon]$$

The following observation directly follows from Definition 14 and Definition 28.

Observation 4. *W_i is a mixture of $\{W_{t_i}\}_{t_i\in\mathcal{T}_i}$ over distribution D_i. For any $\varepsilon > 0$, $W_i^{tr(\varepsilon)}$ is a mixture of $\{W_{t_i}^{tr(\varepsilon)}\}_{t_i\in\mathcal{T}_i}$ over distribution D_i.*

The following observation is useful in later proofs.

Observation 5. *For any $\varepsilon > 0$ and $a' \geq a > 0$, $a \cdot W_i \subseteq a' \cdot W_i$ and $a \cdot W_i^{tr(\varepsilon)} \subseteq a' \cdot W_i^{tr(\varepsilon)}$.*

Proof. Let $c = \frac{a}{a'} \leq 1$. For the first statement, it suffices to prove that $(c\pi_i, cw_i) \in W_i$, for all $(\pi_i, w_i) \in W_i$. Let $\{\sigma_S^{(k)}(t_i)\}_{t_i,S,k}$ be the collection of numbers that satisfy all properties of Definition 14. Then since $c \leq 1$, by considering the collection of numbers $\{c \cdot \sigma_S^{(k)}(t_i)\}_{t_i,S,k}$, we immediately have that $(c\pi_i, cw_i) \in W_i$. For the second statement, let (π'_i, w'_i) be the vector achieved by zeroing out all coordinates (j, t_{ij}) where $f_{ij}(t_{ij}) < \varepsilon$ for the vector (π_i, w_i). By the definition of $W_i^{tr(\varepsilon)}$, we immediately have $(\pi'_i, w'_i) \in W_i^{tr(\varepsilon)}$ and $(c\pi'_i, cw'_i) \in W_i^{tr(\varepsilon)}$. ∎

We next present several desirable properties of the polytopes we consider here.

Lemma 20. *For any $t_i \in \mathcal{T}_i$, any subset of items $B \subseteq [m]$, and any $(x, y) \in W_{t_i}$, consider any $(\hat{x}, \hat{y}) \in [0,1]^{2\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ such that for each $j \in [m]$ and $t'_{ij} \in \mathcal{T}_{ij}$,*

$$\hat{x}(t'_{ij}) = x(t'_{ij})\mathbb{1}[j \in B] \quad \text{and} \quad \hat{y}(t'_{ij}) \leq y(t'_{ij})\mathbb{1}[j \in B].$$

Then $(\widehat{x}, \widehat{y}) \in W_{t_i}$. Moreover, if $(x, y) \in W_{t_i}^{tr(\varepsilon)}$, then $(\hat{x}, \hat{y})$ is also in $W_{t_i}^{tr(\varepsilon)}$. Finally, $W_{t_i}^{tr(\varepsilon)} \subseteq W_{t_i}$.

Proof. It suffices to prove the case where $\hat{y}(t'_{ij}) = y(t'_{ij})\mathbb{1}[j \in B], \forall j, t'_{ij}$, since by Definition 28, we can decrease any $\hat{y}(t'_{ij})$ while maintaining the vector $(\hat{x}, \hat{y})$ to be in W_{t_i}.

Since $(x, y) \in W_{t_i}$, let $\{\sigma_S^{(k)}\}_{S\subseteq[m],k\in[K]}$ be the collection of numbers from Definition 28. Each $\sigma_S^{(k)}$ can be viewed as the probability of the buyer receiving bundle S, and enabling the k-th additive function. Consider another collection of numbers $\{\widehat{\sigma}_S^{(k)}\}_{S\subseteq[m],k\in[K]}$ by simply discarding items in $[m]\backslash B$. Formally, $\widehat{\sigma}_S^{(k)} = \sum_{T\subseteq[m]\backslash B} \sigma_{S\cup T}^{(k)}, \forall S \subseteq B, k \in [K]$, and $\widehat{\sigma}_S^{(k)} = 0$ otherwise. Notice that for every $j \in B$ and $k \in [K]$, $\sum_{S:j\in S} \sigma_S^{(k)} = \sum_{S:j\in S} \widehat{\sigma}_S^{(k)}$. It is not hard to verify that $(\hat{x}, \hat{y})$ and $\{\widehat{\sigma}_S^{(k)}\}_{S,k}$ satisfy all inequalities in Definition 28. Thus $(\hat{x}, \hat{y}) \in W_{t_i}$.

Now $W_{t_i}^{tr(\varepsilon)} \subseteq W_{t_i}$ follows from choosing B to be $\{j : f_{ij}(t_{ij}) \geq \varepsilon\}$. For any $(x, y) \in W_{t_i}^{tr(\varepsilon)}$ and any choice set $B \subseteq [m]$, $(\hat{x}, \hat{y}) \in W_{t_i}$. Since $\hat{x}(t'_{ij}) = \hat{x}(t'_{ij}) \cdot \mathbb{1}[f_{ij}(t_{ij}) \geq \varepsilon]$ and $\hat{y}(t'_{ij}) = \hat{y}(t'_{ij}) \cdot \mathbb{1}[f_{ij}(t_{ij}) \geq \varepsilon]$. By Definition 28, $(\hat{x}, \hat{y})$ also lies in $W_{t_i}^{tr(\varepsilon)}$. □

Lemma 21. *Given any $\varepsilon > 0$ and any distribution $\widetilde{D}_i$ over $\mathcal{T}_i$. Let $\widetilde{W}_i^{tr(\varepsilon)}$ be a mixture of $\left\{W_{t_i}^{tr(\varepsilon)}\right\}_{t_i\in\mathcal{T}_i}$ over $\widetilde{D}_i$, that is, $\widetilde{W}_i^{tr(\varepsilon)} := \sum_{t_i\in\mathcal{T}_i} \Pr_{s\sim\widetilde{D}_i}[s = t_i] \cdot W_{t_i}^{tr(\varepsilon)}$. For each $j \in [m]$, let $S_j \subseteq \mathcal{T}_{ij}$ be any set. For any $(\pi_i, w_i) \in \widetilde{W}_i^{tr(\varepsilon)}$ and any $(\widehat{\pi}_i, \widehat{w}_i) \in [0,1]^{2\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ such that for each $j \in [m]$ and $t'_{ij} \in \mathcal{T}_{ij}$,*

$$\widehat{\pi}_{ij}(t'_{ij}) = \pi_{ij}(t'_{ij})\mathbb{1}[t'_{ij} \in S_j] \quad \text{and} \quad \widehat{w}_{ij}(t'_{ij}) \leq w_{ij}(t'_{ij})\mathbb{1}[t'_{ij} \in S_j].$$

then $(\widehat{\pi}_i, \widehat{w}_i) \in \widetilde{W}_i^{tr(\varepsilon)}$.

Proof. It suffices to prove the case where $\widehat{w}_{ij}(t'_{ij}) = w_{ij}(t'_{ij})\mathbb{1}[t'_{ij} \in S_j], \forall j, t'_{ij}$. This is because $\widetilde{W}_i^{tr(\varepsilon)}$ is a mixture of $\{W_{t_i}^{tr(\varepsilon)}\}_{t_i\in\mathcal{T}_i}$. By Definition 28, for every t_i and any vector $(x, y) \in W_{t_i}^{tr(\varepsilon)}$, we can decrease any $y(t'_{ij})$ while maintaining the vector (x, y) to be in $W_{t_i}^{tr(\varepsilon)}$. Thus, for any $(\widehat{\pi}_i, \widehat{w}_i) \in \widetilde{W}_i^{tr(\varepsilon)}$, it remains in $\widetilde{W}_i^{tr(\varepsilon)}$ after decreasing any $\widehat{w}_{ij}(t'_{ij})$.

For each $t_i \in \mathcal{T}_i$, let $\left(\pi_i^{(t_i)}, w_i^{(t_i)}\right) \in W_{t_i}^{tr(\varepsilon)}$ such that $(\pi_i, w_i) = \sum_{t_i\in\mathcal{T}_i} \Pr_{s\sim\widetilde{D}_i}[s = t_i] \cdot \left(\pi_i^{(t_i)}, w_i^{(t_i)}\right)$. Consider vector $\left(\widehat{\pi}_i^{(t_i)}, \widehat{w}_i^{(t_i)}\right)$ such that for every $j \in [m]$ and $t'_{ij} \in \mathcal{T}_{ij}$

$$\widehat{\pi}_{ij}^{(t_i)}(t'_{ij}) = \pi_{ij}^{(t_i)}(t'_{ij})\mathbb{1}[t_{ij} \in S_j] \quad \text{and} \quad \widehat{w}_{ij}^{(t_i)}(t'_{ij}) = w_{ij}^{(t_i)}(t'_{ij})\mathbb{1}[t_{ij} \in S_j].$$

For each $t_i \in \mathcal{T}_i$, define set $B(t_i) := \{j : t_{ij} \in S_j\}$, by applying Lemma 20 to $\left(\pi_i^{(t_i)}, w_i^{(t_i)}\right)$ and set $B(t_i)$, we have that $\left(\widehat{\pi}_i^{(t_i)}, \widehat{w}_i^{(t_i)}\right) \in W_{t_i}^{tr(\varepsilon)}$.

We notice that by Definition 28, $w_{ij}^{(t_i)}(t'_{ij}) = \pi_{ij}^{(t_i)}(t'_{ij}) = 0$ if $t'_{ij} \neq t_{ij}$. Thus, for every $j \in [m]$ and $t'_{ij} \in \mathcal{T}_{ij}$,

$$\widehat{\pi}_{ij}^{(t_i)}(t'_{ij}) = \pi_{ij}^{(t_i)}(t'_{ij})\mathbb{1}[t'_{ij} \in S_j] \quad \text{and} \quad \widehat{w}_{ij}^{(t_i)}(t'_{ij}) = w_{ij}^{(t_i)}(t'_{ij})\mathbb{1}[t'_{ij} \in S_j].$$

The proof concludes by noticing that $(\widehat{\pi}_i, \widehat{w}_i) = \sum_{t_i\in\mathcal{T}_i} \Pr_{s\sim\widetilde{D}_i}[s = t_i] \cdot \left(\widehat{\pi}_i^{(t_i)}, \widehat{w}_i^{(t_i)}\right) \in \widetilde{W}_i^{tr(\varepsilon)}$. □

The following corollary follows from Observation 4 and Lemma 20.

Corollary 1. *For any* $\varepsilon > 0$, $W_i^{tr(\varepsilon)} \subseteq W_i$.

Proof. This follows easily from the claim that $W_{t_i}^{tr(\varepsilon)} \subseteq W_{t_i}$ (Lemma 20). $\square$

Similar to the constrained-additive case, we define the ε-box polytope of W_i for XOS valuations in Definition 29.

Definition 29. *For* $\varepsilon > 0$*, we denote as* $W_i^{box(\varepsilon)} \subseteq [0,1]^{2\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ *the* ε*-box polytope of* W_i*:* $(\pi_i, w_i) \in W_i^{box(\varepsilon)}$ *if and only if for every* $j \in [m]$ *and* $t_{ij} \in \mathcal{T}_{ij}$ *it holds that*

$$0 \le w_{ij}(t_{ij}) \le \pi_{ij}(t_{ij}) \le \min(\varepsilon, f_{ij}(t_{ij}))$$

The following lemma is similar to Lemma 19.

Lemma 22. *Let* $T = \sum_{i\in[n]}\sum_{j\in[m]}|\mathcal{T}_{ij}|$*, then* $(1-\varepsilon\cdot T)\,W_i^{box(\varepsilon)} \subseteq W_i$*, for all* $\varepsilon < 1/T$.

Proof. For any $(\pi_i, w_i) \in W_i^{box(\varepsilon)}$, we will prove that $(\pi_i, \pi_i) \in W_i$. Then $(\pi_i, w_i) \in W_i$ since $w_{ij}(t_{ij}) \le \pi_{ij}(t_{ij})$ for any j, t_{ij}.

To prove $(\pi_i, \pi_i) \in W_i$, we consider the following set of numbers $\{\sigma_S^{(k)}(t_i)\}_{t_i,S,k}$ (see Definition 14): For each $j \in [m], t_{ij} \in \mathcal{T}_{ij}$, let $c_j(t_{ij}) = \min\left(\frac{\varepsilon}{f_{ij}(t_{ij})}, 1\right)$ and

$$p_j(t_{ij}) = \frac{\pi_{ij}(t_{ij})}{f_{ij}(t_{ij})\cdot\sum_{t_{i,-j}} f_{i,-j}(t_{i,-j})\cdot c_j(t_{ij})\cdot\prod_{j'\ne j}(1-c_{j'}(t_{ij'}))}.$$

Note that for every $j' \in [m]$, there exists a $t_{ij'} \in \mathcal{T}_{ij'}$ such that $f_{ij'}(t_{ij'}) \ge 1/|\mathcal{T}_{ij'}|$. Due to our choice of ε, the corresponding $c_{j'}(t_{ij'}) < 1$. Hence, $\sum_{t_{i,-j}} f_{i,-j}(t_{i,-j})\cdot\prod_{j'\ne j}(1-c_{j'}(t_{ij'})) > 0$, and $p_j(t_{ij})$ is well-defined.

For every t_i, define [26]

$$\sigma_S^{(k)}(t_i) = \begin{cases} p_j(t_{ij})\cdot c_j(t_{ij})\cdot\prod_{j'\ne j}(1-c_{j'}(t_{ij'})), & \text{if } S=\{j\} \text{ and } k = \arg\max_{k'}\alpha_{ij}^{(k')}(t_{ij}) \\ 0, & \text{o.w.}\end{cases}$$

For every j, let $C_j(t_{ij})$ be the independent Bernoulli random variable that activates with probability $c_j(t_{ij})$. Then for every j,

$$\Pr_{C_j(t_{ij}),t_{ij}\sim D_{ij}}[C_j(t_{ij})=1] \le \sum_{t_{ij}\in\mathcal{T}_{ij}} f_{ij}(t_{ij})\cdot\min\left(\frac{\varepsilon}{f_{ij}(t_{ij})},1\right) \le \varepsilon\cdot|\mathcal{T}_{ij}|$$

By the union bound,

$$\sum_{t_{i,-j}} f_{i,-j}(t_{i,-j})\cdot\prod_{j'\ne j}(1-c_{j'}(t_{ij'})) = \Pr_{\substack{t_{i,-j}\sim\mathcal{D}_{i-j}\\ \forall j'\ne j, C_{j'}(t_{ij'})}}[C_{j'}(t_{ij'})=0,\forall j'\ne j] \ge 1-\sum_j \Pr[C_j(t_{ij})=1] \ge 1-\varepsilon\cdot T$$

Now we prove that $(\pi_i, \pi_i) \in W_i$ by verifying all three conditions in Definition 14. For the first condition, since $(\pi_i, w_i) \in (1-\varepsilon T)W_i^{box(\varepsilon)}$, $0 \le w_{ij}(t_{ij}) \le \pi_{ij}(t_{ij}) \le (1-\varepsilon\cdot T)\cdot f_{ij}(t_{ij})\cdot c_j(t_{ij})$. Thus $p_j(t_{ij}) \le 1$ for every j, t_{ij}. We have that $\sum_{S,k}\sigma_S^{(k)}(t_i) = \sum_j p_j(t_{ij})\cdot c_j(t_{ij})\cdot\prod_{j'\ne j}(1-c_{j'}(t_{ij'})) \le \sum_j c_j(t_{ij})\cdot$

[26] When there are two indices $k, k' \in [K]$ such that $k, k^* \in \arg\max_{k'}\alpha_{ij}^{(k')}(t_{ij})$, we break ties in lexicographic order.

$\prod_{j'\neq j}(1-c_{j'}(t_{ij'}))$. We notice that $\sum_j c_j(t_{ij})\cdot\prod_{j'\neq j}(1-c_{j'}(t_{ij'}))$ is exactly the probability that there exists a unique $C_j(t_{ij})=1$. Thus $\sum_{S,k}\sigma_S^{(k)}(t_i)\leq 1$ for all $t_i\in\mathcal{T}_i$.

The second condition:

$$\begin{aligned}&f_{ij}(t_{ij})\cdot\sum_{t_{i,-j}}f_{i,-j}(t_{i,-j})\cdot\sum_{S:j\in S}\sum_{k\in[K]}\sigma_S^{(k)}(t_{ij},t_{i,-j})\\=&f_{ij}(t_{ij})\sum_{t_{i,-j}}f_{i,-j}(t_{i,-j})\cdot p_j(t_{ij})\cdot c_j(t_{ij})\cdot\prod_{j'\neq j}(1-c_{j'}(t_{ij'}))=\pi_{ij}(t_{ij})\end{aligned}$$

The third condition:

$$\begin{aligned}&f_{ij}(t_{ij})\cdot\sum_{t_{i,-j}}f_{i,-j}(t_{i,-j})\cdot\sum_{S:j\in S}\sum_{k\in[K]}\sigma_S^{(k)}(t_{ij},t_{i,-j})\cdot\frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})}\\=&f_{ij}(t_{ij})\sum_{t_{i,-j}}f_{i,-j}(t_{i,-j})\cdot p_j(t_{ij})\cdot c_j(t_{ij})\cdot\prod_{j'\neq j}(1-c_{j'}(t_{ij'}))\cdot\frac{\max_k\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})}=\pi_{ij}(t_{ij})\end{aligned}$$

By Definition 14, $(\pi_i,\pi_i)\in W_i$. Thus, $(\pi_i,w_i)\in W_i$. □

Similar to Lemma 14, we prove in the following lemma that W_i can be approximated by the polytope $\frac{1}{2}W_i^{tr(\varepsilon)}+\frac{1-\varepsilon T}{2}W_i^{box(\varepsilon)}$ within a multiplicative factor.

Lemma 23. *Let $T=\sum_{i\in[n]}\sum_{j\in[m]}|\mathcal{T}_{ij}|$ and any $0\leq\varepsilon<\frac{1}{T}$, then*

$$\frac{1-\varepsilon T}{2}W_i\subseteq\frac{1}{2}W_i^{tr(\varepsilon)}+\frac{1-\varepsilon T}{2}W_i^{box(\varepsilon)}\subseteq W_i.$$

Proof. Let $W'=\frac{1}{2}W_i^{tr(\varepsilon)}+\frac{1-\varepsilon T}{2}W_i^{box(\epsilon)}$. Then $W'\subseteq W_i$ directly follows from Corollary 1 and Lemma 22. We are going to show that for each $(\pi_i,w_i)\in\frac{1-\varepsilon\cdot T}{2}\cdot W_i$, $(\pi_i,w_i)\in W'$. We consider the following vector (π_i^{tr},w_i^{tr}) such that for every $j\in[m],t_{ij}\in\mathcal{T}_{ij}$,

$$\pi_{ij}^{tr}(t_{ij})=\pi_{ij}(t_{ij})\mathbb{1}[f_{ij}(t_{ij})\geq\varepsilon],\qquad w_{ij}^{tr}(t_{ij})=w_{ij}(t_{ij})\mathbb{1}[f_{ij}(t_{ij})\geq\varepsilon]$$

Then by Definition 27 and Observation 5, $(\pi_i^{tr},w_i^{tr})\in\frac{1-\varepsilon\cdot T}{2}W_i^{tr(\varepsilon)}\subseteq\frac{1}{2}W_i^{tr(\varepsilon)}$. Consider the vector (π_i^{box},w_i^{box}) such that

$$\pi_{ij}^{box}(t_{ij})=\pi_{ij}(t_{ij})\mathbb{1}[f_{ij}(t_{ij})<\varepsilon],\qquad w_{ij}^{box}(t_{ij})=w_{ij}(t_{ij})\mathbb{1}[f_{ij}(t_{ij})<\varepsilon]$$

For every $j\in[m],t_{ij}\in\mathcal{T}_{ij}$, since $(\pi_i,w_i)\in\frac{1-\varepsilon\cdot T}{2}\cdot W_i$, by Observation 3 we have $w_{ij}(t_{ij})\leq\pi_{ij}(t_{ij})\leq\frac{1-\varepsilon\cdot T}{2}\cdot f_{ij}(t_{ij})$. Thus $w_{ij}^{box}(t_{ij})\leq\pi_{ij}^{box}(t_{ij})\leq\frac{1-\varepsilon\cdot T}{2}\cdot f_{ij}(t_{ij})\cdot\mathbb{1}[f_{ij}(t_{ij})<\varepsilon]\leq\frac{1-\varepsilon\cdot T}{2}\cdot\min(f_{ij}(t_{ij}),\varepsilon)$. Thus $(\pi_i^{box},w_i^{box})\in\frac{1-\varepsilon T}{2}W_i^{box(\varepsilon)}$. Now observe that $(\pi_i,w_i)=(\pi_i^{tr}+\pi_i^{box},w_i^{tr}+w_i^{box})\in\frac{1}{2}W_i^{tr(\varepsilon)}+\frac{1-\varepsilon T}{2}W_i^{box(\varepsilon)}$, which concludes the proof. □

E.2 Approximating the Single-Bidder Marginal Reduced Form Polytope

In this section we provide the important step that allows us to prove Theorem 11. In the constrained-additive case (Theorem 2), we construct (with high probability) another polytope $\widehat{W_i}$ such that (i) $c\cdot W_i\subseteq\widehat{W_i}\subseteq W_i$ for some absolute constant $c>0$, and (ii) we have an efficient separation oracle for $\widehat{W_i}$. The proof heavily relies on the fact that W_i is down-monotone and thus cannot be easily extended to the single-bidder marginal reduced form polytope W_i in the XOS case. Here we prove a similar statement with a weaker version of property (i): For every vector x in W_i, there exists another vector x' in $\widehat{W_i}$ such that for every coordinate j, $x_j/x'_j\in[a,b]$ for some absolute constant $0<a<b$, and vice versa. It's not hard to see that the original property (i) directly implies this weaker version. A formal statement is shown in Theorem 12.

Theorem 12. *Let $T = \sum_{i\in[n]} \sum_{j\in[m]} |\mathcal{T}_{ij}|$ and b be the bit complexity of the problem instance (Definition 3). For any $i \in [n]$ and $\delta > 0$, there is an algorithm that constructs a convex polytope $\widehat{W}_i \in [0,1]^{2\cdot\sum_{j\in[m]}|\mathcal{T}_{ij}|}$ using poly$(n, m, T, \log(1/\delta))$ samples from D_i, such that with probability at least $1-\delta$ all of the following are satisfied:*

1. *For each $(\widetilde{\pi}_i, \widetilde{w}_i) \in \widehat{W}_i$, there exists a $(\pi_i, w_i) \in W_i$ such that for all $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$:*

$$\frac{\pi_{ij}(t_{ij})}{\widetilde{\pi}_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{3}{2}\right], \quad \frac{w_{ij}(t_{ij})}{\widetilde{w}_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{5}{4}\right]$$

2. *For each $(\pi_i, w_i) \in W_i$, there exists a $(\widetilde{\pi}_i, \widetilde{w}_i) \in \widehat{W}_i$ such that for all $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$:*

$$\frac{\widetilde{\pi}_{ij}(t_{ij})}{\pi_{ij}(t_{ij})} \in \left[\frac{1}{16}, \frac{3}{8}\right], \quad \frac{\widetilde{w}_{ij}(t_{ij})}{w_{ij}(t_{ij})} \in \left[\frac{1}{16}, \frac{5}{16}\right]$$

3. *The vertex-complexity (Definition 10) of $\widehat{W}_i$ is poly$(n, m, T, b, \log(1/\delta))$.*

4. *There exists a separation oracle SO for $\widehat{W}_i$, given access to the value oracle and an adjustable demand oracle (Definition 2) of buyer i's valuation. The running time of SO on any input with bit complexity b' is poly$(n, m, T, b, b', \log(1/\delta))$ and makes poly$(n, m, T, b, b', \log(1/\delta))$ queries to both oracles.*

The algorithm constructs the polytope and the separation orcle SO in time poly$(n, m, T, b, \log(1/\delta))$.

E.2.1 Efficiently Optimizing over the Polytope

Both Corollary 2 and Lemma 24 are useful to show that there exists an efficient separation oracle for our constructed polytope.

In Lemma 24, we show that given access to the adjustable demand oracle and the value oracle, we can optimize any linear objective over the polytope $W_{t_i}^{tr(\varepsilon)}$ (Definition 28).

Lemma 24. *Let $T = \sum_{i\in[n]} \sum_{j\in[m]} |\mathcal{T}_{ij}|$, and b be the bit complexity of the instance. For any $t_i \in \mathcal{T}_i$ and any $\varepsilon > 0$, given access to the adjustable demand oracle $\text{ADEM}_i(\cdot,\cdot,\cdot)$ and the value oracle for buyer i's valuation $v_i(\cdot,\cdot)$, there exists an algorithm that takes arbitrary $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{\sum_j |\mathcal{T}_{ij}|}$ as input and finds*

$$(\pi_i^*, w_i^*) \in \textit{argmax}_{(\pi_i, w_i)\in W_{t_i}^{tr(\varepsilon)}} \boldsymbol{x}\cdot\pi_i + \boldsymbol{y}\cdot w_i.$$

The algorithm runs in time poly$(n, m, T, b, b', 1/\varepsilon)$ and makes $O(m)$ queries to both oracles, where b' is the bit complexity of $(\boldsymbol{x}, \boldsymbol{y})$. Moreover the bit complexity of (π_i^*, w_i^*) is at most $O(bT)$.

Proof. We are going to solve the problem:

$$\begin{aligned} \max \quad & \boldsymbol{x}\cdot\pi_i + \boldsymbol{y}\cdot w_i \\ s.t. \quad & (\pi_i, w_i) \in W_{t_i}^{tr(\varepsilon)} \end{aligned}$$

By Definition 28, $\pi_{ij}(t'_{ij}) = w_{ij}(t'_{ij}) = 0$ if $t'_{ij} \neq t_{ij}$ or $f_{ij}(t_{ij}) < \varepsilon$. Let $Q = \{j \in [m] : f_{ij}(t_{ij}) \geq \varepsilon\}$.

Then according to Definition 28, the problem is equivalent to

$$
\begin{aligned}
(O_1) \quad & \max \sum_{j\in Q} \left(x_j(t_{ij}) \cdot \pi_{ij}(t_{ij}) + y_j(t_{ij}) \cdot w_{ij}(t_{ij})\right) \\
& s.t. \sum_{S\subseteq[m]} \sum_{k\in[K]} \sigma_S^{(k)} \leq 1 \\
& \quad \pi_{ij}(t_{ij}) = \sum_{S:j\in S} \sum_{k\in[K]} \sigma_S^{(k)} && \forall j \in [m] \\
& \quad 0 \leq w_{ij}(t_{ij}) \leq \sum_{S:j\in S} \sum_{k\in[K]} \sigma_S^{(k)} \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})} && \forall j \in [m] \\
& \quad \sigma_S^{(k)} \geq 0 && \forall S \subseteq [m], k \in [K]
\end{aligned}
$$

We notice that at the maximum, $w_{ij}(t_{ij})$ must be equal to $\sum_{S:j\in S}\sum_{k\in[K]} \sigma_S^{(k)} \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})}$ if $y_j(t_{ij}) > 0$, and 0 otherwise. (O_1) is equivalent to (denote $y_j(t_{ij})^+ = \max\{y_j(t_{ij}), 0\}$):

$$
\begin{aligned}
(O_2) \quad & \max \sum_{j\in Q} \left(x_j(t_{ij}) \cdot \pi_{ij}(t_{ij}) + y_j(t_{ij})^+ \cdot w_{ij}(t_{ij})\right) \\
& s.t. \sum_{S\subseteq[m]} \sum_{k\in[K]} \sigma_S^{(k)} \leq 1 \\
& \quad \pi_{ij}(t_{ij}) = \sum_{S:j\in S} \sum_{k\in[K]} \sigma_S^{(k)} && \forall j \in [m] \\
& \quad w_{ij}(t_{ij}) = \sum_{S:j\in S} \sum_{k\in[K]} \sigma_S^{(k)} \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})} && \forall j \in [m] \\
& \quad \sigma_S^{(k)} \geq 0 && \forall S \subseteq [m], k \in [K]
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
(O_3) \quad & \max \sum_{S\subseteq[m]} \sum_{k\in[K]} \sigma_S^{(k)} \left(\sum_{j\in S\cap Q} \left(y_j(t_{ij})^+ \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})} + x_j(t_{ij}) \right) \right) \\
& s.t. \sum_{S\subseteq[m]} \sum_{k\in[K]} \sigma_S^{(k)} \leq 1 \\
& \quad \sigma_S^{(k)} \geq 0 && S \subseteq [m], k \in [K]
\end{aligned}
$$

Clearly, to solve (O_3), it suffices to find $S^* \subseteq Q$ and $k^* \in [K]$ that lies in

$$
\arg\max_{S\subseteq Q, k\in[K]} \sum_{j\in S} \left(y_j(t_{ij})^+ \cdot \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})} + x_j(t_{ij}) \right).
$$

We notice that since $y_j(t_{ij})^+ \cdot \frac{\alpha_{ij}^{(k)}(t_{ij})}{V_{ij}(t_{ij})} \geq 0$ for every j, k, thus $j \in S^*$ for every j such that $x_j(t_{ij}) \geq 0$. Consider the vector $\mathbf{b} \in \mathbb{R}_+^m$ and $\mathbf{p} \in \mathbb{R}_+^m$ such that for each $j \in [m]$,

$$
b_j = \frac{y_j(t_{ij})^+}{V_{ij}(t_{ij})}, \quad p_j = \begin{cases} -x_j(t_{ij})\mathbb{1}[x_j(t_{ij}) < 0], & j \in Q \\ \infty, & j \notin Q \end{cases}
$$

Then $(S^*, k^*) \in \arg\max \sum_{j\in S}\left(b_j\alpha_{ij}^{(k)}(t_{ij}) - p_j\right)$, which can be achieved by a single query to the adjustable demand oracle with input $(t_i, \mathbf{b}, \mathbf{p})$. Now the corresponding vector (π_i^*, w_i^*) can be computed according to (O_2), with $\sigma_S^{(k)} = \mathbb{1}[S = S^* \cup S_+ \wedge k = k^*]$, where $S_+ := \{j \in [m] : x_j(t_{ij}) > 0\}$. [27]

Finally, the bit complexity of each coordinate of (π_i^*, w_i^*) is at most $O(b)$, which implies that the bit complexity of (π_i^*, w_i^*) is at most $O(bT)$. □

Corollary 2. *Let $T = \sum_{i\in[n]}\sum_{j\in[m]} |\mathcal{T}_{ij}|$ and b be the bit complexity of the instance. For any $t_i \in \mathcal{T}_i$, $W_{t_i}^{tr(\varepsilon)}$ has vertex-complexity $O(bT)$.*

E.2.2 Proof of Theorem 12

Choose parameter $\varepsilon = \frac{1}{2T}$ and any $k = \Omega\left(T^4\left(Tb + \log\left(1/\varepsilon\right) + n\log(1/\delta)\right)\right)$ Hence, in this proof, $1 - \varepsilon T = 1/2$. Let $\widehat{D}_i$ be the empirical distribution induced by $N = \left\lceil \frac{32kT}{\varepsilon^2}\right\rceil$ independent samples from D_i. Let $\widehat{W}_i^{tr(\varepsilon)}$ be the mixture of $\{W_{t_i}^{tr(\varepsilon)}\}_{t_i\in T_i}$ over the empirical distribution $\widehat{D}_i$. By Corollary 2, the bit complexity of each corner of polytopes in $\{W_{t_i}^{tr(\varepsilon)}\}_{t_i\in\mathcal{T}_i}$ is at most $O(Tb)$. By Theorem 9, with probability at least $1 - 2Te^{-2Tk} \geq 1-\delta$, both of the following properties hold:

(i) For each $(\widehat{\pi}_i, \widehat{w}_i) \in \widehat{W}_i^{tr(\varepsilon)}$, there exists a $(\pi_i, w_i) \in W_i^{tr(\varepsilon)}$ such that $||(\widehat{\pi}_i, \widehat{w}_i) - (\pi_i, w_i)||_\infty \leq \frac{\varepsilon}{2}$

(ii) For each $(\pi_i, w_i) \in W_i^{tr(\varepsilon)}$, there exists a $(\widehat{\pi}_i, \widehat{w}_i) \in \widehat{W}_i^{tr(\varepsilon)}$ such that $||(\widehat{\pi}_i, \widehat{w}_i) - (\pi_i, w_i)||_\infty \leq \frac{\varepsilon}{2}$

For the rest of the proof, we condition on the event where both properties above hold. Let $W_i' = \frac{1}{2}W_i^{tr(\varepsilon)} + \frac{1}{4}W_i^{box(\varepsilon)}$. Consider the polytope $\widehat{W}_i = \frac{1}{2}\widehat{W}_i^{tr(\varepsilon)} + \frac{1}{4}W_i^{box(\varepsilon)}$. We are going to show that for each $(\pi_i, w_i) \in W_i'$, there exists a $(\widetilde{\pi}_i, \widetilde{w}_i) \in \widehat{W}_i$ such that for all j, t_{ij},

$$\frac{\widetilde{\pi}_{ij}(t_{ij})}{\pi_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{3}{2}\right] \quad \text{and} \quad \frac{\widetilde{w}_{ij}(t_{ij})}{w_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{5}{4}\right].$$

It is not hard to see that this implies Property 2 in the statement of Theorem 12, as $W_i' = \frac{1}{2}W_i^{tr(\varepsilon)} + \frac{1}{4}W_i^{box(\varepsilon)} \supseteq \frac{1}{4}W_i$(due to Lemma 23 and our choice of ε).

For any $(\pi_i, w_i) \in W_i'$, we rewrite (π_i, w_i) as $\frac{1}{2}(\pi_i^{tr}, w_i^{tr}) + \frac{1}{2}(\pi_i^{box}, w_i^{box})$ where $(\pi_i^{tr}, w_i^{tr}) \in W_i^{tr(\varepsilon)}$ and $(\pi_i^{box}, w_i^{box}) \in \frac{1}{2}W_i^{box(\varepsilon)}$. For every $j \in [m]$, we partition $\mathcal{T}_{ij}$ into four disjoint sets:

$$\begin{aligned}
S_j^{(1)} &= \{t_{ij} \in \mathcal{T}_{ij} : f_{ij}(t_{ij}) \geq \varepsilon \wedge \varepsilon < w_{ij}(t_{ij}) \leq \pi_{ij}(t_{ij})\} \\
S_j^{(2)} &= \{t_{ij} \in \mathcal{T}_{ij} : f_{ij}(t_{ij}) \geq \varepsilon \wedge w_{ij}(t_{ij}) \leq \varepsilon < \pi_{ij}(t_{ij})\} \\
S_j^{(3)} &= \{t_{ij} \in \mathcal{T}_{ij} : f_{ij}(t_{ij}) \geq \varepsilon \wedge w_{ij}(t_{ij}) \leq \pi_{ij}(t_{ij}) \leq \varepsilon\} \\
S_j^{(4)} &= \{t_{ij} \in \mathcal{T}_{ij} : f_{ij}(t_{ij}) < \varepsilon\}
\end{aligned}$$

By property (ii), there exists a $(\widehat{\pi}_i^{tr}, \widehat{w}_i^{tr}) \in \widehat{W}_i^{tr(\varepsilon)}$ such that for all $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$:

$$\pi_{ij}^{tr}(t_{ij}) - \varepsilon/2 \leq \widehat{\pi}_{ij}^{tr}(t_{ij}) \leq \pi_{ij}^{tr}(t_{ij}) + \varepsilon/2$$
$$w_{ij}^{tr}(t_{ij}) - \varepsilon/2 \leq \widehat{w}_{ij}^{tr}(t_{ij}) \leq w_{ij}^{tr}(t_{ij}) + \varepsilon/2$$

[27] Both $\alpha_{ij}^{(k^*)}(t_{ij})$ and $V_{ij}(t_{ij}) = \max_k \alpha_{ij}^{(k)}(t_{ij})$ can be computed with $O(1)$ queries to the adjustable demand oracle.

Now consider the following vector $(\widetilde{\pi}_i^{tr}, \widetilde{w}_i^{tr}) \in [0,1]^{2\sum_{j\in[m]}|\mathcal{T}_{ij}|}$:

$$\widetilde{\pi}_{ij}^{tr}(t_{ij}) = \begin{cases} \widehat{\pi}_{ij}^{tr}(t_{ij}) & \text{if } t_{ij} \in S_j^{(1)} \cup S_j^{(2)} \\ 0 & \text{o.w.} \end{cases}$$

$$\widetilde{w}_{ij}^{tr}(t_{ij}) = \begin{cases} \widehat{w}_{ij}^{tr}(t_{ij}) & \text{if } t_{ij} \in S_j^{(1)} \\ 0 & \text{o.w.} \end{cases}$$

In other words,

$$(\widetilde{\pi}_{ij}^{tr}(t_{ij}), \widetilde{w}_{ij}^{tr}(t_{ij})) = (\widehat{\pi}_{ij}^{tr}(t_{ij}) \cdot \mathbb{1}[\pi_{ij}(t_{ij}) > \varepsilon], \widehat{w}_{ij}^{tr}(t_{ij}) \cdot \mathbb{1}[w_{ij}(t_{ij}) > \varepsilon]),$$

and $(\widetilde{\pi}_i^{tr}, \widetilde{w}_i^{tr}) \in \widehat{W}_i^{tr(\varepsilon)}$.

Consider another vector $(\widetilde{\pi}_i^{box}, \widetilde{w}_i^{box}) \in [0,1]^{2\sum_{j\in[m]}|\mathcal{T}_{ij}|}$:

$$(\widetilde{\pi}_{ij}^{box}(t_{ij}), \widetilde{w}_{ij}^{box}(t_{ij})) = \begin{cases} (\pi_{ij}^{box}(t_{ij}), w_{ij}^{box}(t_{ij})) & \text{if } t_{ij} \in S_j^{(1)} \cup S_j^{(4)} \\ (w_{ij}(t_{ij})/2, w_{ij}(t_{ij})/2) & \text{if } t_{ij} \in S_j^{(2)} \\ (\pi_{ij}(t_{ij})/2, w_{ij}(t_{ij})/2) & \text{if } t_{ij} \in S_j^{(3)} \end{cases}$$

For every $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$, clearly $\widetilde{\pi}_{ij}^{box}(t_{ij}) \geq \widetilde{w}_{ij}^{box}(t_{ij})$ since $\pi_{ij}^{box}(t_{ij}) \geq w_{ij}^{box}(t_{ij})$ and $\pi_{ij}(t_{ij}) \geq w_{ij}(t_{ij})$. Moreover, if $t_{ij} \in S_j^{(1)} \cup S_j^{(4)}$, $\widetilde{\pi}_{ij}^{box}(t_{ij}) = \pi_{ij}^{box}(t_{ij}) \leq \frac{1}{2} \cdot \min\{\varepsilon, f_{ij}(t_{ij})\}$, since $(\pi_i^{box}, w_i^{box}) \in \frac{1}{2}W_i^{box(\varepsilon)}$. If $t_{ij} \in S_j^{(2)} \cup S_j^{(3)}$, by the definitions of $S_j^{(2)}$ and $S_j^{(3)}$, we have that $\widetilde{\pi}_{ij}^{box}(t_{ij}) \leq \varepsilon/2 = \frac{1}{2} \cdot \min\{\varepsilon, f_{ij}(t_{ij})\}$. Thus $(\widetilde{\pi}_i^{box}, \widetilde{w}_i^{box}) \in \frac{1}{2}W_i^{box(\varepsilon)}$. Now define $(\widetilde{\pi}_i, \widetilde{w}_i) = \frac{1}{2}(\widetilde{\pi}_i^{tr}, \widetilde{w}_i^{tr}) + \frac{1}{2}(\widetilde{\pi}_i^{box}, \widetilde{w}_i^{box})$. Then $(\widetilde{\pi}_i, \widetilde{w}_i) \in \widehat{W}_i$. It remains to prove that $(\widetilde{\pi}_i, \widetilde{w}_i)$ satisfies: for all j, t_{ij}

$$\frac{\widetilde{\pi}_{ij}(t_{ij})}{\pi_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{3}{2}\right] \quad \text{and} \quad \frac{\widetilde{w}_{ij}(t_{ij})}{w_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{5}{4}\right].$$

We verify all cases based on which set t_{ij} is in.

Case 1: $t_{ij} \in S_j^{(1)}$. Recall that for $t_{ij} \in S_j^{(1)}$, $\varepsilon < w_{ij}(t_{ij}) \leq \pi_{ij}(t_{ij})$. We have that

$$\widetilde{\pi}_{ij}(t_{ij}) - \pi_{ij}(t_{ij}) = \frac{1}{2}(\hat{\pi}_{ij}^{tr}(t_{ij}) - \pi_{ij}^{tr}(t_{ij})).$$

According to property (ii), $\frac{1}{2}(\hat{\pi}_{ij}^{tr}(t_{ij}) - \pi_{ij}^{tr}(t_{ij})) \in [-\frac{\varepsilon}{4}, \frac{\varepsilon}{4}]$, so $\frac{1}{2}(\hat{\pi}_{ij}^{tr}(t_{ij}) - \pi_{ij}^{tr}(t_{ij}))$ also lies in $[-\frac{\pi_{ij}(t_{ij})}{4}, \frac{\pi_{ij}(t_{ij})}{4}]$. Similarly,

$$\widetilde{w}_{ij}(t_{ij}) - w_{ij}(t_{ij}) = \frac{1}{2}(\hat{w}_{ij}^{tr}(t_{ij}) - w_{ij}^{tr}(t_{ij})) \in [-\frac{\varepsilon}{4}, \frac{\varepsilon}{4}] \subseteq [-\frac{w_{ij}(t_{ij})}{4}, \frac{w_{ij}(t_{ij})}{4}].$$

Hence,

$$\frac{\widetilde{\pi}_{ij}(t_{ij})}{\pi_{ij}(t_{ij})} \in \left[\frac{3}{4}, \frac{5}{4}\right] \quad \text{and} \quad \frac{\widetilde{w}_{ij}(t_{ij})}{w_{ij}(t_{ij})} \in \left[\frac{3}{4}, \frac{5}{4}\right].$$

Case 2: $t_{ij} \in S_j^{(2)}$**.** Recall that $\widetilde{w}_{ij}(t_{ij}) = \frac{1}{4} \cdot w_{ij}(t_{ij})$ and $\widetilde{\pi}_{ij}(t_{ij}) = \frac{1}{4} \cdot w_{ij}(t_{ij}) + \frac{1}{2} \cdot \widehat{\pi}_{ij}^{tr}(t_{ij})$. We have that

$$\widetilde{\pi}_{ij}(t_{ij}) \leq \frac{1}{4} \cdot w_{ij}(t_{ij}) + \frac{\pi_{ij}^{tr}(t_{ij})}{2} + \frac{\varepsilon}{4} \leq \frac{3}{2} \cdot \pi_{ij}(t_{ij}).$$

The first inequality follows from property (ii), and the second inequality follows from $\pi_{ij}(t_{ij}) = \frac{\pi_{ij}^{tr}(t_{ij})}{2} + \frac{\pi_{ij}^{box}(t_{ij})}{2} \geq \frac{\pi_{ij}^{tr}(t_{ij})}{2}$ and $w_{ij}(t_{ij}) \leq \varepsilon < \pi_{ij}(t_{ij})$ when $t_{ij} \in S_j^{(2)}$. We also have that

$$\widetilde{\pi}_{ij}(t_{ij}) \geq \frac{\widetilde{\pi}_{ij}^{tr}(t_{ij})}{2} \geq \frac{\pi_{ij}^{tr}(t_{ij})}{2} - \frac{\varepsilon}{4} = \pi_{ij}(t_{ij}) - \frac{\pi_{ij}^{box}(t_{ij})}{2} - \frac{\varepsilon}{4} \geq \pi_{ij}(t_{ij}) - \frac{\varepsilon}{2} \geq \frac{1}{2}\pi_{ij}(t_{ij})$$

The first inequality follows from the non-negativity of $\widetilde{\pi}_{ij}^{box}(t_{ij})$; the second inequality follows from property (ii); the third inequality is due to the fact that $\pi_{ij}^{box}(t_{ij}) \leq \varepsilon/2$; the last inequality is because $\varepsilon < \pi_{ij}(t_{ij})$.

Hence,

$$\frac{\widetilde{\pi}_{ij}(t_{ij})}{\pi_{ij}(t_{ij})} \in \left[\frac{1}{2}, \frac{3}{2}\right] \quad \text{and} \quad \frac{\widetilde{w}_{ij}(t_{ij})}{w_{ij}(t_{ij})} = \frac{1}{4}.$$

Case 3: $t_{ij} \in S_j^{(3)} \cup S_j^{(4)}$**.** Recall that when $t_{ij} \in S_j^{(3)}$, $\widetilde{\pi}_{ij}(t_{ij}) = \frac{1}{4} \cdot \pi_{ij}(t_{ij})$ and $\widetilde{w}_{ij}(t_{ij}) = \frac{1}{4} \cdot w_{ij}(t_{ij})$. When $t_{ij} \in S_j^{(4)}$, as $f_{ij}(t_{ij}) < \varepsilon$, $w_{ij}^{tr}(t_{ij}) = \pi_{ij}^{tr}(t_{ij}) = 0$. Thus $\widetilde{\pi}_{ij}(t_{ij}) = \frac{1}{2}\pi_{ij}^{box}(t_{ij}) = \pi_{ij}(t_{ij})$, and $w_{ij}(t_{ij}) = \frac{1}{2}w_{ij}^{box}(t_{ij}) = w_{ij}(t_{ij})$.

To sum up, we have argued that for each $(\pi_i, w_i) \in W_i'$, there exists a $(\widetilde{\pi}_i, \widetilde{w}_i) \in \widehat{W}_i$ such that for all j, t_{ij},

$$\frac{\widetilde{\pi}_{ij}(t_{ij})}{\pi_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{3}{2}\right] \quad \text{and} \quad \frac{\widetilde{w}_{ij}(t_{ij})}{w_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{5}{4}\right].$$

With a similar analysis, [28] we can also show that for each $(\widetilde{\pi}_i, \widetilde{w}_i) \in \widehat{W}_i$, there exists a $(\pi_i, w_i) \in W_i'$ such that for all j, t_{ij},

$$\frac{\pi_{ij}(t_{ij})}{\widetilde{\pi}_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{3}{2}\right] \quad \text{and} \quad \frac{w_{ij}(t_{ij})}{\widetilde{w}_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{5}{4}\right].$$

Thus Property 1 in the statement of Theorem 12 follows from the fact that $W_i' = \frac{1}{2}W_i^{tr(\varepsilon)} + \frac{1}{4}W_i^{box(\varepsilon)} \subseteq W_i$ (Lemma 23).

Let $\{t_i^{(1)}, ..., t_i^{(N)}\}$ be the N samples from D_i. Then

$$\widehat{W}_i = \sum_{\ell \in [N]} \frac{1}{2N} \cdot W_{t_i^{(\ell)}}^{tr(\varepsilon)} + \frac{1}{4}W_i^{box(\varepsilon)}$$

is the Minkowski addition of $N+1$ polytopes. For Property 3 of the statement, since the vertex-complexity of $W_{t_i}^{tr(\varepsilon)}$ is $O(bT)$ for each t_i (Corollary 2), and the vertex-complexity of $W_i^{box(\varepsilon)}$ is no more than poly(b, T), the vertex-complexity of $\widehat{W}_i$ is no more than $poly(n, m, T, b, \log(1/\delta))$.

At last, we show the existence of an efficient separation oracle SO for $\widehat{W}_i$, by efficiently optimizing any linear objective over $\widehat{W}_i$. Since $\widehat{W}_i$ is the Minkowski addition of polytopes $\{W_{t_i^{(\ell)}}^{tr(\varepsilon)}\}_{\ell \in [N]}$ and $W_i^{box(\varepsilon)}$, in

[28] We only need to switch the role of $(\widetilde{\pi}_i, \widetilde{w}_i)$ and (π_i, w_i). We provide a brief sketch here. First, rewrite $(\widetilde{\pi}_i, \widetilde{w}_i)$ as $\frac{1}{2}(\widehat{\pi}_i^{tr}, \widehat{w}_i^{tr}) + \frac{1}{2}(\widetilde{\pi}_i^{box}, \widetilde{w}_i^{box})$, where $(\widehat{\pi}_i^{tr}, \widehat{w}_i^{tr}) \in \widehat{W}_i^{tr(\varepsilon)}$ and $(\widetilde{\pi}_i^{box}, \widetilde{w}_i^{box}) \in \frac{1}{2}W_i^{box(\varepsilon)}$. Also, redefine $S_j^{(i)}$ in the same fashion but according to $(\widetilde{\pi}_i, \widetilde{w}_i)$. Take $(\pi_i^{tr}, w_i^{tr}) \in W_i^{tr(\varepsilon)}$ to be the point guaranteed to exist by property (i), and define $(\bar{\pi}_{ij}^{tr}(t_{ij}), \bar{w}_{ij}^{tr}(t_{ij})) := (\pi_{ij}^{tr}(t_{ij}) \cdot \mathbb{1}[\widetilde{\pi}_{ij}(t_{ij}) > \varepsilon], w_{ij}^{tr}(t_{ij}) \cdot \mathbb{1}[\widetilde{w}_{ij}(t_{ij}) > \varepsilon])$. Also, define $(\pi_{ij}^{box}(t_{ij}), w_{ij}^{box}(t_{ij}))$ according to which set t_{ij} belongs to in a fashion similar to the proof above. Now define $(\pi_i, w_i) = \frac{1}{2}(\bar{\pi}_i^{tr}, \bar{w}_i^{tr}) + \frac{1}{2}(\pi_i^{box}, w_i^{box})$. Using a similar case analysis, we can prove the claim that $\frac{\pi_{ij}(t_{ij})}{\widetilde{\pi}_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{3}{2}\right]$ and $\frac{w_{ij}(t_{ij})}{\widetilde{w}_{ij}(t_{ij})} \in \left[\frac{1}{4}, \frac{5}{4}\right]$ for all $j \in [m]$ and $t_{ij} \in \mathcal{T}_{ij}$.

order to maximize over $\widehat{W}_i$, it's sufficient to maximize over each polytope. By Lemma 24, we can efficiently optimize any linear objective over $W_{t_i}^{tr(\varepsilon)}$ for every t_i, given the adjustable demand oracle and value oracle. Thus it is sufficient to solve $\max\{\boldsymbol{x}\cdot\pi_i+\boldsymbol{y}\cdot w_i:(\pi_i,w_i)\in W_i^{box(\varepsilon)}\}$ for any vector $\boldsymbol{x},\boldsymbol{y}$. Since in $W_i^{box(\varepsilon)}$, the constraint for each coordinate (j,t_{ij}) is separate: $0\leq w_{ij}(t_{ij})\leq\pi_{ij}(t_{ij})\leq\min(\varepsilon,f_{ij}(t_{ij}))$. Thus the optimum can be achieved by solving the following LP for every coordinate:

$$\begin{aligned}\max\quad & x_j(t_{ij})\cdot\pi_{ij}(t_{ij})+y_j(t_{ij})\cdot w_{ij}(t_{ij})\\ s.t.\quad & 0\leq w_{ij}(t_{ij})\leq\pi_{ij}(t_{ij})\leq\min(\varepsilon,f_{ij}(t_{ij}))\end{aligned}$$

Note that the bit complexity of the output of our optimization algorithm is $poly(n,m,T,b,\log(1/\delta))$. Thus by Theorem 10, there exists a separation oracle SO of $\widehat{W}_i$, that satisfies Property 4 in the statement of Theorem 12.

E.3 Putting Everything Together

In this section, we put all pieces together and provide a complete proof of Theorem 11. Denote (P) the LP in Figure 4 and OPT_{LP} the optimal objective of (P). We consider another LP denoted as (P'). In (P'), in addition to all variables in (P), we introduce new variables $\widehat{\pi}_i=\{\widehat{\pi}_{ij}(t_{ij})\}_{j\in[m],t_{ij}\in\mathcal{T}_{ij}}$ and $\widehat{w}_i=\{\widehat{w}_{ij}(t_{ij})\}_{j\in[m],t_{ij}\in\mathcal{T}_{ij}}$ for every $i\in[n]$. Both (P) and (P') have the same objective function. The only difference between (P) and (P') is that in (P'), we replace Constraint (1) with the following constraints:

$$\text{Constraint }(1'):\quad(\widehat{\pi}_i,\widehat{w}_i)\in\widehat{W}_i,\pi_i\geq\frac{3}{2}\widehat{\pi}_i\geq\mathbf{0},w_i\leq\frac{1}{4}\widehat{w}_i,\quad\forall i\in[n].$$

Here $\widehat{W}_i$ is the proxy polytope from Theorem 12 for each $i\in[n]$. Both inequalities hold coordinate-wisely. Denote OPT'_{LP} the optimal objective of (P'). By Property 4 of Theorem 12, there exists an efficient separation oracle for each $\widehat{W}_i$. Thus we can solve (P') in polynomial time using the Ellipsoid algorithm (Theorem 5). The following lemma shows the relationship between (P) and (P').

Lemma 25. *Suppose for every $i\in[n]$, $\widehat{W}_i$ satisfies the properties in Theorem 12. Then*

- *For any feasible solution $(\pi,w,\hat{\pi},\hat{w},\lambda,\hat{\lambda},\boldsymbol{d})$ to (P'), there exists $\widetilde{\pi}\in[0,1]^{\sum_{i,j}|\mathcal{T}_{ij}|}$ such that $(\widetilde{\pi},w,\lambda,\hat{\lambda},\boldsymbol{d})$ is a feasible solution to (P).*
- $\text{OPT}_{\text{LP}}\leq 64\cdot\text{OPT}'_{\text{LP}}$.

Proof. We prove the first part of the statement. Let $(\pi,w,\hat{\pi},\hat{w},\lambda,\hat{\lambda},\boldsymbol{d})$ be any feasible solution to (P'). Then for every $i\in[n]$, $(\widehat{\pi}_i,\widehat{w}_i)\in\widehat{W}_i$. By Property 1 of Theorem 12, there exists a $(\widetilde{\pi}_i,\widetilde{w}_i)\in W_i$ such that for each $j\in[m]$ and $t_{ij}\in\mathcal{T}_{ij}$,

$$\widetilde{\pi}_{ij}(t_{ij})\leq\frac{3}{2}\widehat{\pi}_{ij}(t_{ij})\leq\pi_{ij}(t_{ij})\quad\widetilde{w}_{ij}(t_{ij})\geq\frac{1}{4}\widehat{w}_{ij}(t_{ij})\geq w_{ij}(t_{ij})$$

Let $\widetilde{\pi}=\{\widetilde{\pi}_i\}_{i\in[n]}$. We are going to show that $(\widetilde{\pi},w,\lambda,\hat{\lambda},\boldsymbol{d})$ is a feasible solution to (P) by verifying all constraints. For Constraint (1), since $(\widetilde{\pi}_i,\widetilde{w}_i)\in W_i$ and $w_{ij}(t_{ij})\leq\widetilde{w}_{ij}(t_{ij}),\forall j,t_{ij}$, thus by Definition 14 $(\widetilde{\pi}_i,w_i)\in W_i$.

For Constraint (2), since $(\pi,w,\hat{\pi},\hat{w},\lambda,\hat{\lambda},\boldsymbol{d})$ is a feasible solution to (P'), for every $j\in[m]$,

$$\sum_i\sum_{t_{ij}\in\mathcal{T}_{ij}}\widetilde{\pi}_{ij}(t_{ij})\leq\sum_i\sum_{t_{ij}\in\mathcal{T}_{ij}}\pi_{ij}(t_{ij})\leq 1.$$

Furthermore, Constraints $(3)-(9)$ are clearly satisfied since $(\pi,w,\hat{\pi},\hat{w},\lambda,\hat{\lambda},\boldsymbol{d})$ is a feasible solution to (P'). Thus $(\widetilde{\pi},w,\lambda,\hat{\lambda},\boldsymbol{d})$ is a feasible solution to (P).

Now we prove the second part of the statement. Let $(\pi^*, w^*, \lambda^*, \hat{\lambda}^*, \boldsymbol{d}^*)$ be any optimal feasible solution to (P). For every buyer i, $(\pi_i^*, w_i^*) \in W_i$. By Property 2 of Theorem 12, there exists a $(\widehat{\pi}_i, \widehat{w}_i) \in \widehat{W}_i$ such that for every j, t_{ij},

$$\widehat{\pi}_{ij}(t_{ij}) \leq \frac{3}{8}\pi_{ij}^*(t_{ij}), \quad \widehat{w}_{ij}(t_{ij}) \geq \frac{1}{16}w_{ij}^*(t_{ij})$$

We are going to show that $\left(\pi = \frac{3}{2}\widehat{\pi}, w = \frac{1}{64}w^*, \hat{\pi}, \hat{w}, \lambda = \frac{1}{64}\lambda^*, \hat{\lambda} = \hat{\lambda}^*, \boldsymbol{d} = \boldsymbol{d}^*\right)$ is a feasible solution to (P'), which implies that $\text{OPT}'_{\text{LP}} \geq \text{OPT}_{\text{LP}}/64$. Firstly, for each $i \in [n]$, $(\widehat{\pi}_i, \widehat{w}_i) \in \widehat{W}_i$ and

$$\pi_i = \frac{3}{2}\widehat{\pi}_i, \qquad w_i = \frac{1}{64}w_i^* \leq \frac{1}{4}\widehat{w}_i$$

Thus Constraint $(1')$ is satisfied. For Constraint (2), since $(\pi^*, w^*, \lambda^*, \hat{\lambda}^*, \boldsymbol{d}^*)$ is a feasible solution to (P), we have that for every j,

$$\sum_{i \in [n]} \sum_{t_{ij} \in \mathcal{T}_{ij}} \pi_{ij}(t_{ij}) = \sum_{i \in [n]} \sum_{t_{ij} \in \mathcal{T}_{ij}} \frac{3}{2}\hat{\pi}_{ij}(t_{ij}) \leq \sum_{i \in [n]} \sum_{t_{ij} \in \mathcal{T}_{ij}} \frac{9}{16}\pi_{ij}^*(t_{ij}) \leq \frac{9}{16} < 1$$

One can easily verify that when $(\pi^*, w^*, \lambda^*, \hat{\lambda}^*, \boldsymbol{d}^*)$ is a feasible solution to (P), $\left(\frac{1}{64}\pi^*, \frac{1}{64}w^*, \frac{1}{64}\lambda^*, \hat{\lambda}^*, \boldsymbol{d}^*\right)$ is also a feasible solution to (P), which implies that $\left(\pi, w, \hat{\pi}, \hat{w}, \lambda, \hat{\lambda}, \boldsymbol{d}\right)$ satisfies Constraints $(3)-(9)$. Thus $\left(\pi, w, \hat{\pi}, \hat{w}, \lambda, \hat{\lambda}, \boldsymbol{d}\right)$ is a feasible solution to (P'). The objective of the solution is a $\frac{1}{64}$-fraction of the objective of $(\pi^*, w^*, \lambda^*, \hat{\lambda}^*, \boldsymbol{d}^*)$ since $\lambda = \frac{1}{64}\lambda^*$, which concludes the proof.

□

Now we are ready to give the proof of Theorem 11.

Proof of Theorem 11:

We consider a fixed $\delta \in (0, 1)$. For each $i \in [n]$, let $\widehat{W}_i$ be the proxy polytope for the single-bidder marginal reduced form polytope W_i that is constructed in Theorem 12 with parameter $\delta' = \frac{\delta}{n}$. Let $\mathcal{E}_1$ be the event that the RPP mechanism computed in Lemma 2 has revenue $\widetilde{\text{PRev}} = \Omega(\text{PRev})$, and let $\mathcal{E}_2$ be the event that for each $i \in [n]$, the proxy polytope $\widehat{W}_i$, satisfies the properties in Theorem 12. By the union bound combined with Lemma 2 and Theorem 12, with probability at least $1 - \delta - \frac{2}{nm}$, both events happen.

We condition on the event that both $\mathcal{E}_1$ and $\mathcal{E}_2$ happens. By Theorem 5 and Property 3 and 4 of Theorem 12, there exists an algorithm that solves the LP (P') in time $\text{poly}(n, m, T, b, \log(1/\delta))$, given access to the adjustable demand oracle and value oracle for all buyers' valuations. Let $(\pi^*, w^*, \hat{\pi}^*, \hat{w}^*, \lambda^*, \hat{\lambda}^*, \boldsymbol{d}^*)$ be an optimal solution to (P'). By Lemma 2, Lemma 25 and Lemma 5, $\text{OPT} \leq 28\text{PRev} + 4\text{OPT}_{LP} \leq \frac{nm}{nm-1}189 \cdot \widetilde{\text{PRev}} + 256 \cdot \text{OPT}'_{LP}$.

Lemma 25 also guarantees the existence of $\widetilde{\pi} \in [0, 1]^{\sum_{i,j}|\mathcal{T}_{ij}|}$ such that $(\widetilde{\pi}, w^*, \lambda^*, \hat{\lambda}^*, \boldsymbol{d}^*)$ is a feasible solution to (P). Although we do not know the value $\widetilde{\pi}$, it turns out sufficient to know λ^* to compute a simple and approximately optimal mechanism. We compute the prices $\{Q_j^*\}_{j \in [m]}$ using λ^* according to Definition 15. In particular,

$$Q_j^* = \frac{1}{2} \cdot \sum_{i \in [n]} \sum_{t_{ij} \in \mathcal{T}_{ij}} f_{ij}(t_{ij}) \cdot V_{ij}(t_{ij}) \cdot \sum_{\substack{\beta_{ij} \in \mathcal{V}_{ij} \\ \delta_{ij} \in \Delta}} \lambda_{ij}^*(t_{ij}, \beta_{ij}, \delta_{ij}) \cdot \mathbb{1}[V_{ij}(t_{ij}) \leq \beta_{ij} + \delta_{ij}],$$

and

$$2 \sum_{j \in [m]} Q_j^* = \text{OPT}'_{LP}.$$

According to Theorem 6, we can construct a two-part tariff mechanism $\mathcal{M}_{\text{TPT}}$ with prices $\{Q_j^*\}_{j\in[m]}$ and a rationed posted price mechanism $\mathcal{M}_{\text{PP}}$ (computed in Lemma 2) in time $\text{poly}(n,m,T)$, such that

$$\text{OPT}'_{LP} = 2\sum_{j\in[m]} Q_j^* = O(\text{Rev}(\mathcal{M}_{\text{TPT}})) + O(\widetilde{\text{PRev}}).$$

To sum up, we can compute in time $\text{poly}(n,m,T,b,\log(1/\delta))$ a two-part tariff mechanism $\mathcal{M}_{\text{TPT}}$ and a rationed posted price mechanism $\mathcal{M}_{\text{PP}}$, such that $\text{OPT} \le c_1 \cdot \text{Rev}(\mathcal{M}_{\text{PP}}) + c_2 \cdot \text{Rev}(\mathcal{M}_{\text{TPT}})$ for some absolute constants $c_1, c_2 > 0$ with probability at least $1-\delta-\frac{2}{nm}$. □

F Missing Details from Section 4

Definition 30. *The Kolmogorov distance between two distributions $\mathcal{D}$ and $\widehat{\mathcal{D}}$ supported on $\mathbb{R}$ is defined as*

$$d_K(\mathcal{D},\widehat{\mathcal{D}}) = \sup_{z\in\mathbb{R}} \left| \Pr_{t\sim\mathcal{D}}[t\le z] - \Pr_{\widehat{t}\sim\widehat{\mathcal{D}}}[\widehat{t}\le z] \right|$$

We need the following robustness result from Cai and Daskalakis [CD17].

Theorem 13 (Theorem 3 in [CD17])**.** *Suppose all bidders' valuations are constrained additive. Let $\mathcal{M}$ be a Sequential Posted Price with Entry Fee Mechanism (as defined in Mechanism 3 [29]) whose prices are $\{p_{ij}\}_{i\in[n],j\in[m]}$ and its entry fee function for the i-th bidder is $\xi_i : 2^{[m]} \to \mathbb{R}_+$. If for each $i\in[n]$ and $j\in[m]$ we have $d_K(\mathcal{D}_{ij},\widehat{\mathcal{D}}_{ij}) \le \varepsilon$, and $\mathcal{D}_{i,j}$ and $\widehat{\mathcal{D}}_{i,j}$ are both supported on $[0,1]$ (that is when each bidder's value for a single item is at most 1), then*

$$|\text{Rev}(\mathcal{M},\mathcal{D}) - \text{Rev}(\mathcal{M},\widehat{\mathcal{D}})| \le 4nm^2\varepsilon,$$

where $\text{Rev}(\mathcal{M},\mathcal{D})$ and $\text{Rev}(\mathcal{M},\widehat{\mathcal{D}})$ are the revenues of $\mathcal{M}$ under $\mathcal{D} = \times_{i,j}\mathcal{D}_{ij}$ and $\widehat{\mathcal{D}} = \times_{i,j}\widehat{\mathcal{D}}_{ij}$ respectively.

Mechanism 3 Sequential Posted Price with Entry Fee Mechanism (SPEM)

0: Before the mechanism starts, the seller decides on a collection of $\{p_{ij}\}_{i\in[n],j\in[m]}$ and a collection of entry fee functions $\{\xi_i(\cdot)\}_{i\in[n]}$, where $\xi_i : 2^{[m]} \to \mathbb{R}_+$ is buyer i's entry fee.
1: Bidders arrive sequentially in the lexicographical order.
2: When buyer i arrive, the seller shows her the set of available items $S\subseteq[m]$, as well as their prices $\{p_{ij}\}_{j\in S}$ and asks buyer i to pay an *entry fee* of $\xi_i(S)$. Note that S is the set of items that are not purchased by the first $i-1$ buyers.
3: **if** Bidder i pays the entry fee $\delta_i(S)$ **then**
4: $\quad i$ receives her favorite bundle S_i^* and pays $\sum_{j\in S_i^*} p_{ij}$.
5: $\quad S \leftarrow S\backslash S_i^*$.
6: **else**
7: $\quad i$ gets nothing and pays 0.
8: **end if**

Lemma 26. *Suppose all bidders' valuations are constrained additive. If for each $i\in[n]$ and $j\in[m]$ we have $d_K(\mathcal{D}_{ij},\widehat{\mathcal{D}}_{ij}) \le \varepsilon$, and $\mathcal{D}_{i,j}$ and $\widehat{\mathcal{D}}_{i,j}$ are both supported on $[0,1]$ (that is when each bidder's value for a single item is at most 1), then*

$$c_1\cdot\text{OPT}(\widehat{\mathcal{D}}) - O(nm^2\varepsilon) \le \text{OPT}(\mathcal{D}) \le \frac{\text{OPT}(\widehat{\mathcal{D}})}{c_2} + O(nm^2\varepsilon)$$

[29] Indeed the result holds for any buyers' order, we choose the lexicographical order to keep the notation light.

for some absolute constant $c_1, c_2 > 0$, where $\text{OPT}(\mathcal{D})$ (or $\text{OPT}(\widehat{\mathcal{D}})$) is the optimal revenue for distribution $\mathcal{D} = \times_{i,j} \mathcal{D}_{ij}$ (or $\widehat{\mathcal{D}} = \times_{i,j} \widehat{\mathcal{D}}_{ij}$).

Proof. Let M_1, M_2 be the optimal RPP and TPT for $\mathcal{D}$, and we denote their revenue as $\text{Rev}(M_1)$, $\text{Rev}(M_2)$ respectively. Let M_3, M_4 be the optimal RPP and TPT for $\widehat{\mathcal{D}}$, and we denote their revenue as $\text{Rev}(M_3)$, $\text{Rev}(M_4)$ respectively. Since any TPT is a SPEM, by Theorem 13, we know that $|\text{Rev}(M_2) - \text{Rev}(M_4)| \leq 4nm^2\varepsilon$, as

$$\text{Rev}(M_4) \geq \text{Rev}(M_2, \widehat{\mathcal{D}}) \geq \text{Rev}(M_2, \mathcal{D}) - 4nm^2\varepsilon = \text{Rev}(M_2) - 4nm^2\varepsilon,$$

and

$$\text{Rev}(M_2) \geq \text{Rev}(M_4, \mathcal{D}) \geq \text{Rev}(M_4, \widehat{\mathcal{D}}) - 4nm^2\varepsilon = \text{Rev}(M_4) - 4nm^2\varepsilon.$$

Similarly, since any RPP is also a SPEM if we treat the buyers' valuation as Unit-Demand, we have $|\text{Rev}(M_1) - \text{Rev}(M_3)| \leq 4nm^2\varepsilon$. By Lemma 1, $\max\{\text{Rev}(M_1), \text{Rev}(M_2)\} \geq \Omega(\text{OPT}(\mathcal{D}))$. Hence, $\text{OPT}(\widehat{\mathcal{D}}) \geq \max\{\text{Rev}(M_3), \text{Rev}(M_4)\} \geq \max\{\text{Rev}(M_1), \text{Rev}(M_2)\} - 4nm^2\varepsilon \geq \Omega(\text{OPT}(\mathcal{D})) - 4nm^2\varepsilon$. Similarly, $\text{OPT}(\mathcal{D}) \geq \Omega(\text{OPT}(\widehat{\mathcal{D}})) - 4nm^2\varepsilon$.

□

Proof of Theorem 4: For each D_{ij}, we first take $O\left(\frac{\log(nm/\delta)}{\varepsilon^2}\right)$ samples. Let $\widehat{D}_{ij}$ be the uniform distribution over the samples. By the union bound and the DKW inequality [DKW56], $d_K(D_{ij}, \widehat{D}_{ij}) \leq \varepsilon$ for all $i \in [n]$ and $j \in [m]$ with probability at least $1 - \delta$. Now we run the algorithm from Theorem 1 on $\times_{i,j} \widehat{D}_{ij}$ and let $\mathcal{M}$ be the computed mechanism. Note that $\text{Rev}(\mathcal{M}, \widehat{D})$, the revenue of $\mathcal{M}$ under $\widehat{D} = \times_{i,j} \widehat{D}_{ij}$, is $\Omega(\text{OPT}(\widehat{D}))$ – the optimal revenue for distribution $\widehat{D}$. By Theorem 13, $\text{Rev}(\mathcal{M}, D) \geq \text{Rev}(\mathcal{M}, \widehat{D}) - O(nm^2\varepsilon)$. By Lemma 26, $\text{OPT}(\widehat{D}) \geq \Omega(\text{OPT}(D)) - O(nm^2\varepsilon)$. Chaining the ineuqalities together, we have that $\text{Rev}(\mathcal{M}, D) \geq \Omega(\text{OPT}(D)) - O(nm^2\varepsilon)$. □

G Counterexample for Adjustable Demand Oracle

For XOS valuations, our algorithm for constructing the simple mechanism requires access to a special adjustable demand oracle $\text{ADem}_i(\cdot, \cdot, \cdot)$. Readers may wonder if this enhanced oracle (rather than a demand oracle) is necessary to prove our result. In this section we show that (even an approximation of) ADem_i can not be implemented using polynomial number of queries from the value oracle, demand oracle and a classic XOS oracle. All the oracles are defined as follows. Throughout this section, we only consider a single buyer and thus drop the subscript i. Recall that the XOS valuation $v(\cdot)$ satisfies that $v(S) = \max_{k \in [K]} \left\{\sum_{j \in S} \alpha_j^{(k)}\right\}$ for every set S, where $\{\alpha_j^{(k)}\}_{j \in [m]}$ is the k-th additive function.

- Demand Oracle (Dem): takes a price vector $p \in \mathbb{R}^m$ as input, and outputs $S^* \in \arg\max_{S \subseteq [m]} \left(v(S) - \sum_{j \in S} p_j\right)$.
- XOS Oracle (Xos): takes a set $S \subseteq [m]$ as input, and outputs the k^*-th additive function $\{\alpha_j^{(k^*)}\}_{j \in [m]}$, where $k^* \in \arg\max_{k \in [K]} \left\{\sum_{j \in S} \alpha_j^{(k)}\right\}$.
- Value Oracle: takes a set $S \subseteq [m]$ as input, and outputs $v(S)$. We notice that a value oracle can be easily simulated with an XOS oracle. Thus we focus on XOS oracles for the rest of this section.
- Adjustable Demand Oracle (ADem): takes a coefficient vector $b \in \mathbb{R}^m$ and a price vector $p \in \mathbb{R}^m$ as inputs, and outputs $(S^*, \{\alpha_j^{(k^*)}\}_{j \in [m]})$ where $(S^*, k^*) \in \arg\max_{S \subseteq [m], k \in [K]} \left\{\sum_{j \in S} b_j \alpha_j^{(k)} - \sum_{j \in S} p_j\right\}$.

An (approximate) implementation of ADEM is an algorithm that takes inputs $b, p \in \mathbb{R}^m$, and outputs a set $S \subseteq [m]$ and $k \in [K]$. The algorithm has access to the demand oracle and XOS oracle of v. We denote $\text{ALG}(v, b, p)$ the output of the algorithm. For any $\alpha > 1$, ALG is an α-approximation to ADEM if for every XOS valuation v and every $b, p \in \mathbb{R}^m$, the algorithm outputs (S', k') that satisfies:

$$\max_{S\subseteq[m],k\in[K]} \left\{\sum_{j\in S} b_j\alpha_j^{(k)} - \sum_{j\in S} p_j\right\} \leq \alpha \cdot \left(\sum_{j\in S} b_j\alpha_j^{(k')} - \sum_{j\in S'} p_j\right)$$

In Theorem 14 we show that we cannot approximate the output of an Adjustable Demand Oracle within any finite factor, if we are permitted to query polynomial many times the XOS, Value and Demand Oracle.

Theorem 14. *Given any $\alpha > 1$, there **does not exist** an implementation of* ADEM *(denoted as* ALG*) that satisfies both of the following properties:*

1. *For any XOS valuation v over m items,* ALG *makes poly(m) queries to the value oracle, the demand oracle and XOS oracle of v, and runs in time poly(m, b). Here b is the bit complexity of the problem instance (See Definition 3).*

2. ALG *is an α-approximation to* ADEM.

Proof. Recall that a value oracle can be easily simulated with an XOS oracle. Thus we only argue for demand queries and XOS queries. For the sake of contradiction, assume there exists such an algorithm ALG. Let $\ell > e^{2\alpha}$ be an arbitrary even integer. Let $m = 2\ell$. Denote by $L = \binom{\ell}{\ell/2}$. We decompose the items into two sets $S_1 = \{1, \ldots, \ell\}$ and $S_2 = \{\ell + 1, \ldots, m\}$.

We consider an XOS valuation $\widehat{v}$ generated by 2ℓ additive functions denoted by $\{\widehat{\alpha}_j^{(k)}\}_{j\in[m],k\in[2\ell]}$ parameterized by variables $\varepsilon', \varepsilon$ such that $0 < (\ell + 1)\varepsilon' < \varepsilon < \frac{1}{2}$. For $k = 1, \ldots, \ell$ and $j \in [m]$:

$$\widehat{\alpha}_j^{(k)} = \begin{cases} j + \varepsilon + \left(1 - \frac{1}{\ell}\right)\varepsilon' & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

For $k = \ell + 1, \ldots, 2\ell$ and $j \in [m]$, define

$$\widehat{\alpha}_j^{(k)} = \begin{cases} j & \text{if } j = k - \ell \\ 0 & \text{if } j \in S_1 \text{ and } j \neq k - \ell \\ \frac{2}{\ell} \cdot \varepsilon & \text{if } j \in S_2 \end{cases}$$

Next, we introduce a family of XOS valuations $\{v_r\}_{r\in[L]}$ over m items. For every r, the valuation v_r is generated by $K = 2\ell + 1$ additive functions, denoted as $\{\alpha_{r,j}^{(k)}\}_{j\in[m]}, \forall k \in [K]$. We define all the additive functions as follows (and hence v_r is defined as $v_r(S) = \max_{k\in[K]} \sum_{j\in S} \alpha_{r,j}^{(k)}, \forall S \subseteq [m]$):

For $k \in [2\ell]$: For every $r \in [L]$ and $j \in [m]$, define

$$\alpha_{r,j}^{(k)} = \widehat{\alpha}_j^{(k)}$$

For $k = 2\ell + 1$: Take any bijective mapping $\mathcal{C}$ between $[L]$ and subsets of S_2 of size $\frac{\ell}{2}$. For every $r \in [L]$, define

$$\alpha_{r,j}^{(2\ell+1)} = \begin{cases} 1 & \text{if } j \in S_1 \text{ and } j \neq \ell \\ 1+\varepsilon & \text{if } j = \ell \\ \frac{2}{\ell} \cdot \varepsilon' & \text{if } j \in \mathcal{C}(r) \\ 0 & \text{if } j \in S_2 \backslash \mathcal{C}(r) \end{cases}$$

In the following lemmas, we prove that given access to both the XOS and Demand oracle, the algorithm can not distinguish between valuations v_r and $\widehat{v}$, unless the algorithm knows $\mathcal{C}(r)$ (or r).

Lemma 27. *For any $r \in [L]$ and any nonempty set $S \subseteq [m]$, $\sum_{j\in S} \alpha_{r,j}^{(2\ell+1)} > \sum_{j \in S} \alpha_{r,j}^{(k)}$ for all $k \in [2\ell]$ if and only if $S = S_1 \cup \mathcal{C}(r)$. Hence for any $r \in [\ell]$ and any nonempty set $S \subseteq [m]$,*

$$\text{XOS}(v_r, S) = \begin{cases} \{\alpha_{r,j}^{(2\ell+1)}\}_{j\in[m]} & \textit{if } S = S_1 \cup \mathcal{C}(r) \\ \text{XOS}(\widehat{v}, S) & \textit{otherwise} \end{cases}$$

Proof. Let $j^* = \max\{j : j \in S_1 \cap S\}$ (define it to be 0 if $S \cap S_1 = \emptyset$). We consider the following cases:

- Case 1: $S \cap S_1 = \emptyset$. Then $S \subseteq S_2$. For any $k^* = \ell+1, \ldots, 2\ell$,

$$\sum_{j\in S} \alpha_{r,j}^{(k^*)} = \sum_{j \in S} \frac{2}{\ell}\varepsilon > \sum_{j\in S}\frac{2}{\ell}\varepsilon' \geq \sum_{j \in S\cap\mathcal{C}(r)} \frac{2}{\ell}\varepsilon' = \sum_{j\in S}\alpha_{r,j}^{(2\ell+1)}$$

- Case 2: $S \cap S_2 = \emptyset$. Then $S \subseteq S_1$ and $S \cap S_1 \neq \emptyset$, which implies that $j^* > 0$. We have

$$\sum_{j\in S}\alpha_{r,j}^{(j^*)} = j^* + \varepsilon + \left(1 - \frac{1}{\ell}\right)\varepsilon' > j^* + \varepsilon \geq \sum_{j=1}^{j^*} \alpha_{r,j}^{(2\ell+1)} \geq \sum_{j\in S}\alpha_{r,j}^{(2\ell+1)}$$

  Here the second last inequality follows from the fact that $S \subseteq S_1$ and $\alpha_{r,j}^{(2\ell+1)} = 1$ for every $j \in S_1/\{\ell\}$ and $\alpha_{r,\ell}^{(2\ell+1)} = 1+\varepsilon$. The last inequality follows from the definition of j^*.

- Case 3: $S_1 \not\subseteq S$, $S \cap S_1 \neq \emptyset$ and $S \cap S_2 \neq \emptyset$. We have that

$$\sum_{j\in S}\alpha_{r,j}^{(\ell+j^*)} = j^* + \sum_{j\in S_2\cap S}\frac{2}{\ell}\varepsilon > j^* + \sum_{j \in S_2 \cap S}\frac{2}{\ell}\varepsilon' \geq \sum_{j\in S}\alpha_{r,j}^{(2\ell+1)}$$

  The last inequality follows from the fact that $\sum_{j\in S\cap S_1}\alpha_{r,j}^{(2\ell+1)} \leq j^*$: If $j^* \neq \ell$, then $\alpha_{r,j}^{(2\ell+1)} = 1, \forall j \in S \cap S_1$; If $j^* = \ell$, since $S_1 \not\subseteq S$, $\sum_{j \in S\cap S_1}\alpha_{r,j}^{(2\ell+1)} = |S \cap S_1| + \varepsilon \leq (\ell - 1) + \varepsilon < \ell$.

- Case 4: $S_1 \subseteq S$ and $S \cap S_2 \neq \mathcal{C}(r)$. We have

$$\sum_{j\in S}\alpha_{r,j}^{(\ell)} = \alpha_{r,\ell}^{(\ell)} = \ell + \varepsilon + \left(1 - \frac{1}{\ell}\right)\varepsilon' > (\ell+\varepsilon) + \sum_{j\in\mathcal{C}(r)\cap S}\frac{2}{\ell}\varepsilon' = \sum_{j \in S}\alpha_{r,j}^{(2\ell+1)}$$

  Here the inequality is because: $\mathcal{C}(r) \cap S \neq \mathcal{C}(r)$, then $|\mathcal{C}(r) \cap S| \leq \frac{\ell}{2} - 1$, which implies that $\sum_{j\in\mathcal{C}(r)\cap S}\frac{2}{\ell}\varepsilon' \leq \left(\frac{\ell}{2} - 1\right)\frac{2}{\ell}\varepsilon' < \left(1 - \frac{1}{\ell}\right)\varepsilon'$.

- Case 5: $S_1 \subseteq S$ and $S \cap S_2 = \mathcal{C}(r)$. Recall that $|\mathcal{C}(r)| = \frac{\ell}{2}$. We notice that $\sum_{j \in S} \alpha_{r,j}^{(2\ell+1)} = \ell + \varepsilon + \varepsilon'$. On the other hand,

$$\max_{k \in [\ell]} \sum_{j \in S} \alpha_{r,j}^{(k)} = \sum_{j \in S} \alpha_{r,j}^{(\ell)} = \ell + \varepsilon + \left(1 - \frac{1}{\ell}\right) \varepsilon' < \sum_{j \in S} \alpha_{r,j}^{(2\ell+1)}$$

and

$$\max_{\ell < k \leq 2\ell} \sum_{j \in S} \alpha_{r,j}^{(k)} = \sum_{j \in S} \alpha_{r,j}^{(2\ell)} = \ell + \sum_{j \in \mathcal{C}(r)} \alpha_{r,j}^{(2\ell)} = \ell + \varepsilon < \sum_{j \in S} \alpha_{r,j}^{(2\ell+1)}$$

We have proved the first part of the statement. The second part of the statement then follows by noticing that the first 2ℓ additive functions of v_r and $\widehat{v}$ are exactly the same. □

Lemma 28. *For any $r \in [L]$ and any set of prices $\boldsymbol{p} \in \mathbb{R}^m_{\geq 0}$ such that $\{j \in S_2 : p_j \leq \frac{2}{\ell}\varepsilon'\} \neq \mathcal{C}(r)$, we have $\textsc{Dem}(v_r, \boldsymbol{p}) = \textsc{Dem}(\widehat{v}, \boldsymbol{p})$.*

Proof. Recall that

$$\textsc{Dem}(v_r, \boldsymbol{p}) \in \arg\max_{S \subseteq [m]} \left\{ v_r(S) - \sum_{j \in S} p_j \right\} = \arg\max_{S \subseteq [m]} \left\{ \max_{k \in [K]} \sum_{j \in S} (\alpha_{r,j}^{(k)} - p_j) \right\}$$

We notice that valuations v_r and $\widehat{v}$ differ only in the $(2\ell + 1)$-th additive valuation (with coefficients $\{\alpha_{r,j}^{(2\ell+1)}\}_{j \in [m]}$). Thus by Lemma 27, $\textsc{Dem}(v_r, \boldsymbol{p}) = \textsc{Dem}(\widehat{v}, \boldsymbol{p})$ unless $S^* = S_1 \cup \mathcal{C}(r)$ is the favorite bundle for valuation v_r at price $\boldsymbol{p}$, i.e.

$$\sum_{j \in S_1 \cup \mathcal{C}(r)} (\alpha_{r,j}^{(2\ell+1)} - p_j) = \max_{S \subseteq [m], k \in [K]} \sum_{j \in S} (\alpha_{r,j}^{(k)} - p_j)$$

Let $S_0 = \{j \in S_2 : p_j \leq \frac{2}{\ell}\varepsilon'\}$. Firstly, if $\mathcal{C}(r) \not\subseteq S_0$, then there exists $j \in S_2$ such that $p_j > \frac{2}{\ell}\varepsilon'$. Since $\alpha_{r,j}^{(2\ell+1)} = \frac{2}{\ell}\varepsilon', \forall j \in \mathcal{C}(r)$, we have

$$\sum_{j \in S_1 \cup \mathcal{C}(r)} (\alpha_{r,j}^{(2\ell+1)} - p_j) < \sum_{j \in S_1 \cup (\mathcal{C}(r) \cap S_0)} (\alpha_{r,j}^{(2\ell+1)} - p_j) \leq \max_{S \subseteq [m], k \in [K]} \sum_{j \in S} (\alpha_{r,j}^{(k)} - p_j),$$

which implies that $\textsc{Dem}(v_r, \boldsymbol{p}) = \textsc{Dem}(\widehat{v}, \boldsymbol{p})$. It remains to consider the case where $\mathcal{C}(r) \subsetneq S_0$ and $\mathcal{C}(r) \neq S_0$. We have

$$\sum_{j \in S_1 \cup \mathcal{C}(r)} (\alpha_{r,j}^{(2\ell+1)} - p_j) \leq \ell + \varepsilon - \sum_{j \in S_1} p_j + \sum_{j \in \mathcal{C}(r)} \alpha_{r,j}^{(2\ell+1)} = \ell + \varepsilon - \sum_{j \in S_1} p_j + \varepsilon'$$

Here the first inequality follows from $\sum_{j \in S_1} \alpha_{r,j}^{(2\ell+1)} = \ell + \varepsilon$ and $p_j \geq 0, \forall j \in \mathcal{C}(r)$. And the equality follows from $\sum_{j \in \mathcal{C}(r)} \alpha_{r,j}^{(2\ell+1)} = |\mathcal{C}(r)| \cdot \frac{2}{\ell}\varepsilon' = \ell/2 \cdot \frac{2}{\ell}\varepsilon' = \varepsilon'$. On the other hand,

$$\begin{aligned}
\sum_{j\in S_1\cup S_0}(\alpha_{r,j}^{(2\ell)}-p_j) &= \ell-\sum_{j\in S_1}p_j+\sum_{j\in S_0}\left(\alpha_{r,j}^{(2\ell)}-p_j\right)\\
&\geq \ell-\sum_{j\in S_1}p_j+\sum_{j\in S_0}\left(\frac{2}{\ell}\varepsilon-\frac{2}{\ell}\varepsilon'\right)\\
&\geq \ell-\sum_{j\in S_1}p_j+\left(\frac{\ell}{2}+1\right)\left(\frac{2}{\ell}\varepsilon-\frac{2}{\ell}\varepsilon'\right)\\
&= \ell-\sum_{j\in S_1}p_j+\varepsilon-\varepsilon'+\frac{2}{\ell}\left(\varepsilon-\varepsilon'\right)\\
&> \ell-\sum_{j\in S_1}p_j+\varepsilon-\varepsilon'+\frac{2}{\ell}\left((\ell+1)\varepsilon'-\varepsilon'\right)\\
&= \ell+\varepsilon+\varepsilon'-\sum_{j\in S_1}p_j
\end{aligned}$$

Here the first inequality follows from $p_j \leq \frac{2}{\ell}\varepsilon'$, $\alpha_{r,j}^{(2\ell)} = \frac{2}{\ell}\varepsilon$ for $j \in S_0$. The second inequality follows from $|S_0| \geq \frac{\ell}{2}+1$, since $|\mathcal{C}(r)| = \frac{\ell}{2}$, $\mathcal{C}(r) \subseteq S_0$, and $\mathcal{C}(r) \neq S_0$. The third inequality follows from our choice of ε and ε' such that $\varepsilon > (\ell+1)\varepsilon'$. Thus,

$$\sum_{j\in S_1\cup\mathcal{C}(r)}(\alpha_{r,j}^{(2\ell+1)}-p_j) < \sum_{j\in S_1\cup S_0}(\alpha_{r,j}^{(2\ell)}-p_j) \leq \max_{S\subseteq[m],k\in[K]}\sum_{j\in S}(\alpha_{r,j}^{(k)}-p_j)$$

which implies that $\textsc{Dem}(v_r,\boldsymbol{p}) = \textsc{Dem}(\widehat{v},\boldsymbol{p})$. □

To complete the proof of Theorem 14, set $\varepsilon = 0.1$ and $\varepsilon' = \frac{0.1}{2(\ell+1)}$. We notice that the bit complexity of our input is $b = O(\text{poly}(\ell))$ for any valuation v_r. Now consider the coefficient vector $\boldsymbol{c} = (1, 1/2, \ldots, 1/\ell, 0, \ldots, 0)$ with price vector $\boldsymbol{p} = \boldsymbol{0}$. For any valuation v_r in the family, clearly ADEM will select the whole set $[m]$ since all coefficients are non-negative.

For any $k \in [2\ell]$, $\sum_{j\in[m]} c_j\alpha_{r,j}^{(k)} \leq 1+\varepsilon+\left(1-\frac{1}{\ell}\right)\varepsilon' < 2$. On the other hand, when $k = 2\ell+1$, since $\ell > e^{2\alpha}$ we have

$$\sum_{j\in[m]} c_j\alpha_{r,j}^{(2\ell+1)} \geq \sum_{j\in[\ell]}\frac{1}{j}\cdot 1 > \log(\ell) > 2\alpha > \alpha\cdot\max_{k\neq 2\ell+1}\left\{\sum_{j\in[m]}c_j\alpha_{r,j}^{(k)}\right\} = \alpha\cdot\max_{k}\left\{\sum_{j\in[m]}c_j\widehat{\alpha}_j^{(k)}\right\}$$

Thus informally speaking, to obtain an α-approximation to ADEM for every valuation v_r, the algorithm needs to distinguish v_r from $\widehat{v}$ by identifying the $(2\ell+1)$-th additive function. By Lemmas 27 and 28, the algorithm must be able to identify r or $\mathcal{C}(r)$. However, there are $L = \binom{\ell}{\ell/2}$ different valuations in the family $\{v_r\}_{r\in[L]}$. Thus there must exist one v_r such that the algorithm can not distinguish v_r from $\widehat{v}$ within $poly(\ell) = poly(m,b)$ queries to both oracles.

Formally, consider the execution of ALG on valuation $\widehat{v}$. Let Q be the total number of queries during the execution and define set $S^{(1)},\ldots,S^{(Q)} \subseteq [m]$ as follows: For every $q \in [Q]$, if the q-th query is an XOS query, define $S^{(q)}$ as the input set to the XOS oracle; If it's a demand query, let $S^{(q)} = \{j \in S_2 : p_j^{(q)} \leq \frac{2}{\ell}\varepsilon'\}$, where $\boldsymbol{p}^{(q)} = \{p_j^{(q)}\}_{j\in[m]}$ is the input price vector to the demand oracle. Recall that $m = 2\ell$ and $b = \text{poly}(\ell)$, we have $Q = \text{poly}(m,b) < L = \binom{\ell}{\ell/2}$ for sufficiently large ℓ. Thus there exists some $r^* \in [L]$ such that $\mathcal{C}(r^*) \neq S_2 \cap S^{(q)}$ for any $q \in [Q]$. By Lemma 27, we have that $\textsc{Xos}(\widehat{v},S^{(q)}) = \textsc{Xos}(v_r,S^{(q)})$ and by

Lemma 28 we have that $\textsc{Dem}(\widehat{v}, \boldsymbol{p}^{(q)}) = \textsc{Dem}(v_r, \boldsymbol{p}^{(q)})$ for any $q \in [Q]$. This implies that the execution (and thus output) of ALG on input valuation $\widehat{v}$ is exactly the same as its execution on v_{r^*}.

We notice that from the above calculation, $\max_k \left\{\sum_{j\in[m]} c_j \alpha_{r^*,j}^{(k)}\right\} \geq \sum_{j\in[\ell]} \frac{1}{j} > \log(\ell) > 2\alpha$, while $\max_k \left\{\sum_{j\in[m]} c_j \widehat{\alpha}_j^{(k)}\right\} = \max_{k\neq 2\ell+1} \left\{\sum_{j\in[m]} c_j \alpha_{r^*,j}^{(k)}\right\} < 2$. Thus on input $(v_{r^*}, \boldsymbol{c})$, ALG achieves less than a $\frac{1}{\alpha}$-fraction of the optimal objective from ADEM, contradicting with the fact that ALG is an α-approximation to ADEM. □

References

[AFH+12] Saeed Alaei, Hu Fu, Nima Haghpanah, Jason Hartline, and Azarakhsh Malekian. Bayesian Optimal Auctions via Multi- to Single-agent Reduction. In *the 13th ACM Conference on Electronic Commerce (EC)*, 2012.

[Ala11] Saeed Alaei. Bayesian Combinatorial Auctions: Expanding Single Buyer Mechanisms to Many Buyers. In *the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2011.

[BCD20] Johannes Brustle, Yang Cai, and Constantinos Daskalakis. Multi-item mechanisms without item-independence: Learnability via robustness. EC '20, page 715–761, New York, NY, USA, 2020. Association for Computing Machinery.

[BILW14] Moshe Babaioff, Nicole Immorlica, Brendan Lucier, and S. Matthew Weinberg. A Simple and Approximately Optimal Mechanism for an Additive Buyer. In *the 55th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2014.

[CD11] Yang Cai and Constantinos Daskalakis. Extreme-Value Theorems for Optimal Multidimensional Pricing. In *the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2011.

[CD17] Yang Cai and Constantinos Daskalakis. Learning multi-item auctions with (or without) samples. In *the 58th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2017.

[CD21] Yang Cai and Constantinos Daskalakis. Recommender systems meet mechanism design. *arXiv preprint arXiv:2110.12558*, 2021.

[CDO+15] Xi Chen, Ilias Diakonikolas, Anthi Orfanou, Dimitris Paparas, Xiaorui Sun, and Mihalis Yannakakis. On the complexity of optimal lottery pricing and randomized mechanisms. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 1464–1479. IEEE, 2015.

[CDW12a] Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. An Algorithmic Characterization of Multi-Dimensional Mechanisms. In *the 44th Annual ACM Symposium on Theory of Computing (STOC)*, 2012.

[CDW12b] Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Optimal Multi-Dimensional Mechanism Design: Reducing Revenue to Welfare Maximization. In *the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2012.

[CDW12c] Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Optimal Multi-Dimensional Mechanism Design: Reducing Revenue to Welfare Maximization. *arxiv report*, 2012. http://arxiv.org/abs/1207.5518.

[CDW13a] Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Reducing Revenue to Welfare Maximization : Approximation Algorithms and other Generalizations. In *the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2013.

[CDW13b] Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Understanding Incentives: Mechanism Design becomes Algorithm Design. In *the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2013.

[CDW16] Yang Cai, Nikhil R. Devanur, and S. Matthew Weinberg. A duality based unified approach to bayesian mechanism design. In *the 48th Annual ACM Symposium on Theory of Computing (STOC)*, 2016.

[CH13] Yang Cai and Zhiyi Huang. Simple and Nearly Optimal Multi-Item Auctions. In *the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2013.

[CHK07] Shuchi Chawla, Jason D. Hartline, and Robert D. Kleinberg. Algorithmic Pricing via Virtual Valuations. In *the 8th ACM Conference on Electronic Commerce (EC)*, 2007.

[CHMS10a] Shuchi Chawla, Jason Hartline, David Malec, and Balasubramanian Sivan. Sequential posted pricing and multi-parameter mechanism design. *CoRR*, abs/0907.2435, 2010.

[CHMS10b] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-Parameter Mechanism Design and Sequential Posted Pricing. In *the 42nd ACM Symposium on Theory of Computing (STOC)*, 2010.

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

[CMS15] Shuchi Chawla, David L. Malec, and Balasubramanian Sivan. The power of randomness in bayesian optimal mechanism design. *Games and Economic Behavior*, 91:297–317, 2015.

[CO21] Yang Cai and Argyris Oikonomou. On simple mechanisms for dependent items. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, EC '21, page 242–262, New York, NY, USA, 2021. Association for Computing Machinery.

[COVZ21] Yang Cai, Argyris Oikonomou, Grigoris Velegkas, and Mingfei Zhao. An efficient ε-bic to bic transformation and its application to black-box reduction in revenue maximization. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1337–1356. SIAM, 2021.

[CWD21] Yang Cai, S. Matthew Weinberg, and Constantinos Daskalakis. Reducing mechanism design to algorithm design. *Manuscript*, 2021.

[CZ16] Yang Cai and Mingfei Zhao. Simple mechanisms for subadditive buyers via duality. *CoRR*, abs/1611.06910, 2016.

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

[DDK18] Constantinos Daskalakis, Nishanth Dikkala, and Gautam Kamath. Testing ising models. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1989–2007. Society for Industrial and Applied Mathematics, 2018.

[DDT14] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. The complexity of optimal mechanism design. In *the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2014.

[DKL20] Paul Dütting, Thomas Kesselheim, and Brendan 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.

[DKW56] Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, pages 642–669, 1956.

[DW12] Constantinos Daskalakis and S. Matthew Weinberg. Symmetries and Optimal Multi-Dimensional Mechanism Design. In *the 13th ACM Conference on Electronic Commerce (EC)*, 2012.

[GLS81] Martin Grötschel, László Lovász, and Alexander Schrijver. The Ellipsoid Method and its Consequences in Combinatorial Optimization. *Combinatorica*, 1(2):169–197, 1981.

[GLS12] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012.

[HN12] Sergiu Hart and Noam Nisan. Approximate Revenue Maximization with Multiple Items. In *the 13th ACM Conference on Electronic Commerce (EC)*, 2012.

[KP80] Richard M. Karp and Christos H. Papadimitriou. On Linear Characterizations of Combinatorial Optimization Problems. In *the 21st Annual Symposium on Foundations of Computer Science (FOCS)*, 1980.

[KSM+19] Pravesh Kothari, Sahil Singla, Divyarthi Mohan, Ariel Schvartzman, and S Matthew Weinberg. Approximation schemes for a unit-demand buyer with independent items via symmetries. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 220–232. IEEE, 2019.

[KW12] Robert Kleinberg and S. Matthew Weinberg. Matroid Prophet Inequalities. In *the 44th Annual ACM Symposium on Theory of Computing (STOC)*, 2012.

[Mye81] Roger B. Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[Rub16] Aviad Rubinstein. On the computational complexity of optimal simple mechanisms. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 21–28, 2016.

[RW15] Aviad Rubinstein and S. Matthew Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, pages 377–394, 2015.

[Yao15] Andrew Chi-Chih Yao. An n-to-1 bidder reduction for multi-item auctions and its applications. In *SODA*, 2015.