On Simple Mechanisms for Dependent Items

Yang Cai*
Yale University, USA
yang.cai@yale.edu

Argyris Oikonomou†
Yale University, USA
argyris.oikonomou@yale.edu

Abstract

We study the problem of selling n heterogeneous items to a single buyer, whose values for different items are *dependent*. Under arbitrary dependence, Hart and Nisan [30] show that no simple mechanism can achieve a non-negligible fraction of the optimal revenue even with only two items. We consider the setting where the buyer's type is drawn from a correlated distribution that can be captured by a Markov Random Field, one of the most prominent frameworks for modeling high-dimensional distributions with structure.

If the buyer's valuation is additive or unit-demand, we extend the result to all MRFs and show that $\max\{\text{SRev}, \text{BRev}\}$ can achieve an $\Omega\left(\frac{1}{e^{O(\Delta)}}\right)$-fraction of the optimal revenue, where Δ is a parameter of the MRF that is determined by how much the value of an item can be influenced by the values of the other items. We further show that the exponential dependence on Δ is unavoidable for our approach and a polynomial dependence on Δ is unavoidable for any approach. When the buyer has a XOS valuation, we show that $\max\{\text{SRev}, \text{BRev}\}$ achieves at least an $\Omega\left(\frac{1}{e^{O(\Delta)}+\frac{1}{\sqrt{n\gamma}}}\right)$-fraction of the optimal revenue, where γ is the spectral gap of the Glauber dynamics of the MRF. Note that in the special case of independently distributed items, $\Delta = 0$ and $\frac{1}{n\gamma} \leq 1$, and our results recover the known constant factor approximations for a XOS buyer [41]. We further extend our parametric approximation to several other well-studied dependency measures such as the *Dobrushin coefficient* [27] and the *inverse temperature*. In particular, we show that if the MRF is in the *high temperature regime*, $\max\{\text{SRev}, \text{BRev}\}$ is still a constant factor approximation to the optimal revenue even for a XOS buyer. Our results are based on the Duality-Framework by Cai et al. [14] and a new concentration inequality for XOS functions over dependent random variables.

*Supported by a Sloan Foundation Research Fellowship and the NSF Award CCF-1942583 (CAREER).
†Supported by a Sloan Foundation Research Fellowship.

1 Introduction

The design of revenue-optimal auctions for selling multiple items is a central problem in Economics and Computer Science. In the past decade, significant progress has been made, first in efficient computation of revenue-optimal auctions [18, 19, 1, 10, 2, 11, 12, 15, 13, 3, 6, 24], and then in the identification of *simple auctions* that achieve constant factor approximations to the optimal revenue [4, 45, 41, 14, 21, 17] under the *item-independence* assumptions. [1] Despite being theoretically appealing, item-independence is an unrealistic assumption in practice. In this paper, we go beyond the item-independence assumption and study simple and approximately optimal auctions for selling *dependent items*.

Unfortunately, strong negative results exist if we allow the items to be arbitrarily dependent [8, 30]. For example, Hart and Nisan [30] show that the revenue of the best deterministic mechanism is unboundedly smaller than the revenue of the optimal randomized mechanism even when we are only selling two correlated items to a single buyer. Since all simple mechanisms in the literature are deterministic, the result also implies that no simple mechanism that has been considered so far can provide any guarantee to the revenue for even two correlated items. Arguably, however, high-dimensional distributions that arise in practice are rarely arbitrary, as arbitrary high-dimensional distributions cannot be represented efficiently, and are known to require exponentially many samples to learn or even perform the most basic statistical tests on them; see e.g. [25] for a discussion. To overcome the curse of dimensionality, a major focus of Statistics and Machine Learning has been on identifying and exploiting the structural properties of high-dimensional distributions for succinct representation, efficient learning, and efficient statistical inference. There are several widely-studied frameworks to model the structure of dependence in high-dimensional distributions. In this work, we propose capturing the dependence between item values using one of the most prominent graphical models – *Markov Random Fields (MRFs)*. Note that MRFs are fully general and can be used to express arbitrary high-dimensional distributions. The main advantage of MRFs is that there are several natural complexity parameters that allow the user to tune the dependence structure in the distributions represented by MRFs from product measures all the way up to arbitrary distributions. **Our goal is to provide parametric approximation ratios of simple mechanisms that degrade gracefully with respect to these natural parameters.**

MRFs are formally defined in Definition 2. Intuitively, a MRF can be thought of as a graph (or a hypergraph) where each node represents a random variable (or item value in our case). There is a potential function associated with each edge that captures the correlation between the two incident random variables. How does it represent a joint distribution? The probability for a particular realization or the random variables, or known as a configuration of the random field, is proportional to the exponential of the total potential of the configuration. MRF is a flexible model. For example, we can capture the degree of (positive or negative) correlations between two random variables by controlling the corresponding potential function. Here we provide a stylistic example to illustrate the suitability of MRFs for modeling buyers' joint value distributions. Imagine that we manage a car dealership. A potential buyer is hoping to purchase one car, i.e., has a unit-demand valuation. The dealership carries various brands and types of vehicles, and will like to find the optimal way to price each car. However, it would be naïve to assume the buyer's value for each car is independently distributed. The example in Figure 1 demonstrates how a MRF can better capture the customer's joint value distribution for different cars.

[1] Intuitively, item-independence means that each bidder's value for each item is independently distributed, and this definition has been suitably generalized to set value functions such as submodular or subadditive functions [41].

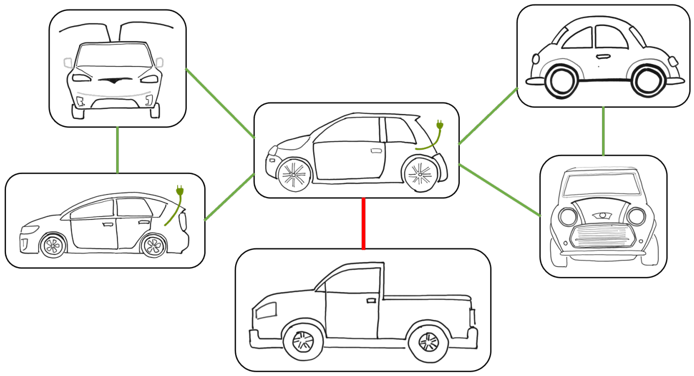

Figure 1: We draw a green edge (or a red edge) between two cars if their values are positively correlated (or negative correlated). The car in the center is an electric coupe with a retro design. Its value is positively correlated with the two electric cars on its left and the two small coupes on its right, but is negatively correlated with the pick-up truck.

1.1 Main Results and Techniques

We focus on the single buyer case and allow the buyer's valuation to be as general as a XOS function (a.k.a. a fractionally subadditive function). [2] We consider the two most extensively studied forms of simple mechanisms: *selling the items separately* and *selling the grand bundle*. We use SRev and BRev to denote the optimal revenue obtainable by these two types of mechanisms respectively. In a sequence of papers, it was shown that $\max\{\text{SRev}, \text{BRev}\}$ is a constant factor approximation to the optimal revenue for a single additive or unit-demand buyer under the item-independence assumption [18, 4, 14]. [3] Our first main result extends the above approximation to any MRFs. The approximation ratio degrades with the *maximum weighted degree* Δ that captures the degree of dependence among the item values.

Parameter I: Maximum Weighted Degree Δ. The formal definition can be found in Definition 4. As we mentioned, a MRF can be thought of as a graph (or a hypergraph) where each node represents a random variable. The weight of an edge is related to the maximum absolute value the corresponding potential function can take and represents the "strength" of the dependence between two incident random variables. The weighted degree of a random variable is simply the sum of weights from all incident edges. If the maximum weighted degree Δ of a MRF is small, then no random variable can depend strongly on many other random variables. Note that $\Delta = 0$ when the random variables are independent, and the instance constructed by Hart and Nisan [30] corresponds to a MRF with $\Delta = \infty$.

Result I: For a single additive or unit-demand buyer whose type is generated by a MRF with maximum weighted degree Δ, $\max\{\text{SRev}, \text{BRev}\} = \Omega\left(\frac{\text{OPT}}{\exp(O(\Delta))}\right)$, where OPT is the optimal revenue.

The formal statement of the result is in Theorem 1 and 2. We further show that the dependence on Δ is necessary. For any sufficiently large number C, there exists a MRF with $\Delta = O(C)$ such

[2] The class of XOS functions is a super-class of submodular functions, and is contained in the class of subadditive functions.

[3] More specifically, SRev denotes the optimal expected revenue achievable by any posted price mechanism. When the buyer has a unit-demand valuation, [18, 14] show that SRev is already a constant factor approximation of the optimal revenue.

that $\max\{\text{SRev}, \text{BRev}\}$ is no more than $\frac{\text{OPT}}{C^{1/7}}$ (Theorem 3) using a modification of the Hart-Nisan construction [30]. Although there is still an exponential gap between our upper and lower bounds, it shows that whenever Result I fails to provide a constant factor approximation (independent of the number of items), no constant factor approximation is possible without further restrictions on the dependency. We leave it as an open question to close the gap between our upper and lower bounds. The main tool we use is a generalization of the prophet inequality to the case where the rewards are sampled from a MRF (Lemma 3). The overall analysis is similar to the one used by Cai et al. [14] for the item-independent case. We show that the exponential dependence on Δ is unavoidable for this type of analysis in Theorem 8. More specifically, a key step of the analysis involves approximating the optimal revenue in a single-dimensional setting, known as the copies setting, using SRev. Theorem 8 constructs an instance such that the optimal revenue in the copies setting is at least $\exp(\Delta)$ times larger than $\max\{\text{SRev}, \text{BRev}\}$.

Rubinstein and Weinberg [41] show that, under the item-independence assumption, $\max\{\text{SRev}, \text{BRev}\}$ is still a constant factor approximation to the optimal revenue for a buyer whose valuation is a subadditive function. Our second main result extends their result to any MRFs when the buyer's valuation is a XOS function. The approximation ratio depends on Δ and the *spectral gap of the Glauber Dynamics* γ.

Parameter II: Spectral Gap of the Glauber Dynamics γ. A common way to generate a sample from a high-dimensional distribution is via a Markov Chain Monte Carlo method known as the Glauber dynamics (see Definition 5). The spectral gap γ of the Glauber dynamics is the difference between the largest eigenvalue $\lambda_1 = 1$ and the second largest eigenvalue λ_2 of the transition matrix of the Glauber dynamics. It is well-known that λ_2 is strictly less than 1 for any MRFs [36], so γ is always strictly positive.

Result II: For a single XOS buyer whose type is generated by a MRF, $\max\{\text{SRev}, \text{BRev}\} = \Omega\left(\frac{\text{OPT}}{\exp(O(\Delta)) + \frac{1}{\sqrt{n\gamma}}}\right)$, where n is the number of items, γ is the spectral gap of the Glauber Dynamics, and Δ is the maximum weighted degree. [a]

[a] Although the approximation ratio depends on n, the ratio indeed improves if we increase n and fix γ .

Some remarks are in order. First, our approximation ratio holds for any MRFs. Second, for any n-dimensional random vector $X = (x_1, \ldots, x_n)$, the X_i's are considered *weakly dependent* if the spectral gap $\gamma = \Omega(\frac{1}{n})$. For example, when the x_i's are independent, $\gamma \geq \frac{1}{n}$. Finally, the condition $\gamma \geq \Omega(\frac{1}{n})$ is extensively studied in probability theory. The condition is satisfied under the Dobrushin uniqueness condition (see Section 6 for details), a well-known sufficient condition that ensures weak dependency; it implies rapid mixing of the Glauber dynamics (i.e., they mix in time $O(n \log n)$); it also guarantees that polynomial functions concentrate in Ising models [29, 26].

The formal statement of Result II can be found in Theorem 4. The analysis follows the same general framework by Cai and Zhao [17]. The major new challenge is to prove that any XOS function $g(X)$ concentrates, when X is a drawn from a high-dimensional distribution D. Proving concentration inequalities for non-linear functions over dependent random variables is a non-trivial task that lies at the heart of many high-dimensional statistical problems. We prove a parametric concentration inequality for XOS functions that depends on the spectral gap of the Glauber dynamics for D (Lemma 13). The proof is based on a combination of the Poincaré inequality and a special property of XOS functions – the self-boundingness. We believe this concentration inequality may be of independent interest. An interesting question is whether the approximation ratio needs to depend on both Δ and γ. We show that the dependence on Δ is crucial, as no approximation can be obtained with only restriction on the spectral gap even for a single additive or unit-demand buyer (Theorem 7). [4] We do not know whether it is possible to obtain an approximation that only depends on Δ for a XOS buyer and leave it as an open question. [5] We suspect such an improvement requires proving a parametric concentration inequality for XOS functions that only depend on the maximum weighted degree Δ, which we believe will have further applications.

Our Results under Other Weak Dependence Conditions. There are several alternative ways to parametrize the degree of dependency in a high-dimensional distribution. We focus on two prominent ones – the *Dobrushin coefficient* and the *inverse temperature of a MRF*, and discuss how our approximation results change under these conditions. We first consider the Dobrushin coefficient and its relaxations. An important concept is the influence matrix.

Influence Matrix and the Dobrushin Condition For any n-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_n)$, we define the influence of variable j on variable i as

$$\alpha_{i,j} := \sup_{\substack{x_{-i-j} \\ x_j \neq x'_j}} d_{TV}\left(F_{X_i|X_j=x_j, X_{-i-j}=x_{-i-j}}, F_{X_i|X_j=x'_j, X_{-i-j}=x_{-i-j}}\right), \text{ [6]}$$

where $F_{X_i|X_{-i}=x_{-i}}$ denotes the conditional distribution of X_i given $X_{-i} = x_{-i}$. Let $\alpha_{i,i} := 0$ for each i. We define the influence matrix $A := (\alpha_{i,j})_{i,j\in[n]}$. When the X_i's are weakly dependent, the entries of A should have small values. The *Dobrushin Coefficient*, defined as $||A||_\infty = \max_{i\in[n]} \sum_{j\in[n]} \alpha_{i,j}$, was originally introduced by Dobrushin [27] in the study of Gibbs measures. The Dobrushin coefficient less than 1 is known as the *Dobrushin uniqueness condition*, under which the Gibbs distribution has a unique equilibrium, hence the name. The condition can be viewed as a sufficient condition that guarantees weak dependence and has been extensively studied in statistical physics and probability literature (see e.g. [28, 43]). As the spectral radius of any matrix is no more than its L_∞ norm, a relaxation of the Dobrushin uniqueness condition is to restrict the spectral radius ρ of A to be less than 1. We show that $n\gamma \geq 1-\rho$ (Lemma 16), so we can replace the dependence on $n\gamma$ with $1-\rho$ in Result II when the item values are weakly dependent (Theorem 5). We also show that the dependence on Δ is necessary. Without any restriction on Δ, the gap between $\max\{\text{SRev}, \text{BRev}\}$ and the optimal revenue could be unbounded even under the Dobrushin uniqueness condition for an additive or unit-demand buyer (Theorem 7). Next, we consider how the approximation guarantee degrades in terms of the inverse temperature of a MRF.

Inverse Temperature β of a MRF The inverse temperature is related to both the maximum weighted degree and the Dobrushin coefficient. See Definition 3 for the formal definition. Intuitively, as the inverse temperature increases (or temperature drops), the dependence between the different random variables strengthens. When the inverse temperature is 0, the MRF represents a product distribution. The *high temperature* regime is when the inverse temperature is less than 1. This parameter often controls when phase transitions in the behavior of MRFs happen, and hence the name. The Dobrushin coefficient always upper bounds the inverse temperature. Recently, MRFs in the high temperature regime have been applied to model weakly dependent random variables [22].

[4]Indeed, we prove an even stronger result that shows no finite approximation ratio is possible under only the Dobrushin uniqueness condition, which implies that $\gamma = \Omega(\frac{1}{n})$ (Lemma 16).

[5]A naïve approach is to directly bound γ using a function of Δ. However, this approach can at best provide an approximation ratio that is exponential in n, as $\frac{1}{\gamma}$ could be exponential in n even when Δ is upper bounded by some absolute constant [39].

[6]$d_{TV}(\cdot,\cdot)$ denotes the total variation distance between two distributions, hence $\alpha_{i,j}$ measures the maximum total variation distance we can have between two conditional distributions of variable i that only differ on the value of variable j.

	Maximum Weighted Degree Δ	Maximum Weighted Degree Δ and Spectral Gap γ	Spectral Gap γ or Dobrushin Coefficient $\alpha < 1$	Inverse Temperature $\beta < 1$
Additive or Unit-Demand	UB: $\Omega\left(\frac{\text{OPT}}{\exp(O(\Delta))}\right)$ (Theorem 2 and 1) LB: $O\left(\frac{\text{OPT}}{\Delta^{1/7}}\right)$ (Theorem 3)	UB: $\Omega\left(\frac{\text{OPT}}{\exp(O(\Delta))}\right)$ ($\leftarrow$) LB: $O\left(\frac{\text{OPT}}{\Delta^{1/7}}\right)$ ($\leftarrow$)	Unbounded (Theorem 7)	UB: $\Omega\left(\sqrt{1-\beta}\cdot\text{OPT}\right)$ ($\downarrow$) LB: open
XOS	UB: open LB: $O\left(\frac{\text{OPT}}{\Delta^{1/7}}\right)$ ($\uparrow$)	UB: $\Omega\left(\frac{\text{OPT}}{\exp(O(\Delta))+\frac{1}{\sqrt{n\gamma}}}\right)$ (Theorem 4) LB: $O\left(\frac{\text{OPT}}{\Delta^{1/7}}\right)$ ($\leftarrow$)	Unbounded ($\uparrow$)	UB: $\Omega\left(\sqrt{1-\beta}\cdot\text{OPT}\right)$ (Theorem 6) LB: open

Table 1: The table contains our upper bounds and lower bounds of the approximation ratio of $\max\{\text{SRev}, \text{BRev}\}$ in various settings. The results are listed based on (i) the valuation of the buyer and (ii) the parameters the approximation ratio can depend on. In our table, an arrow means the result follows from the result that the arrow points to.

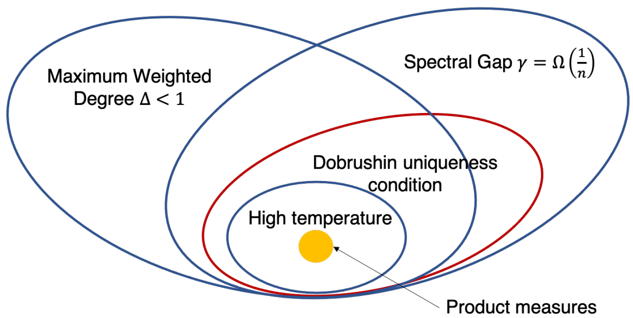

Figure 2: The relationship between the parameters: inverse temperature, Dobrushin coefficient, maximum weighted degree, and spectral gap of the Glauber dynamics.

We show that if the MRF is in the high temperature regime, then its maximum weighted degree $\Delta < 1$ and the spectral gap γ of the Glauber dynamics has value at least $\frac{1-\beta}{n}$. As a corollary of Result II, we have

Result III: For a single XOS buyer, $\max\{\text{SRev}, \text{BRev}\} = \Omega\left(\sqrt{1-\beta}\cdot\text{OPT}\right)$, where $\beta < 1$ is the inverse temperature.

The result states that as long as the inverse temperature is bounded away from 1 by any constant, $\max\{\text{SRev}, \text{BRev}\}$ achieves a constant fraction of the optimal revenue. Theorem 6 contains the formal statement of the result.

We summarize our results in Table 1 and the relationship between the parameters in Figure 2.

1.2 Related Work

Simple vs. Optimal Auctions There has been a large body of work on multi-item auction design focusing on either approximation results under item-independence [18, 19, 1, 15, 37, 4, 45, 41, 14, 21, 17] or impossibility to approximate under arbitrary dependence [8, 30]. Two types of models have been studied for items with limited dependence. The first model considers a specific type of dependence where each item's value is a linear combination of "independent features" [20, 5]. Unlike MRFs, this model cannot express arbitrary structure of dependence. Indeed, the values of any two items can only be positively correlated under this model. The second model considers the smoothed complexity of the problem [38]. Their result applies to arbitrary dependence structure between the item values, but only achieves an approximation ratio that is exponential in the number of items. Our paper is the first to consider a model general enough to capture arbitrary structure of dependence and obtain parametric approximation ratios that are independent of the number of items.

MRFs and Weakly Dependent Random Variables There has been growing interest in understanding the behavior of weakly dependent random variables that can be captured by a MRF in the high temperature regime or under the Dobrushin uniqueness condition [29, 26, 22]. In mechanism design, Brustle et al. [9] is the first to propose modeling dependent item values using MRFs in multi-item auctions, but they focus on the sample complexity of learning nearly optimal auctions.

2 Preliminaries

Basic Notation We consider an auction where a seller is selling n heterogeneous items to a single buyer. We denote the buyer's type $\boldsymbol{t}$ as $\langle t_i \rangle_{i=1}^n$, where t_i is the buyer's private information about item i. We use D to denote the distribution of $\boldsymbol{t}$, D_i to denote the marginal distribution of t_i, and $D_{i|c_{-i}}$ to denote the distribution of t_i conditioned on $t_{-i} = c_{-i}$. We use $\textsc{Supp}(\mathcal{F})$ to denote the support of distribution $\mathcal{F}$, and $T_i = \textsc{Supp}(D_i)$ and $T = \textsc{Supp}(D)$. Moreover, we use $f(c)$ to denote $\Pr_{\boldsymbol{t}\sim D}[\boldsymbol{t} = c]$. For any item i and any $c_i \in T_i$ and $c_{-i} \in T_{-i}$, we use $f_i(c_i)$ to denote $\Pr_{t_i \sim D_i}[t_i = c_i]$, $f_i(c_i \mid c_{-i})$ to denote $\frac{\Pr_{\boldsymbol{t}\sim D}[t=(c_i,c_{-i})]}{\Pr_{\boldsymbol{t}\sim D}[t_{-i}=c_{-i}]}$, $f_{-i}(c_{-i})$ to denote $\Pr_{\boldsymbol{t}\sim D}[t_{-i} = c_{-i}]$, and $f_{-i}(c_{-i} \mid c_i)$ to denote $\frac{\Pr_{\boldsymbol{t}\sim D}[t=(c_i,c_{-i})]}{\Pr_{\boldsymbol{t}\sim D}[t_i=c_i]}$. We also define $F_i(c_i) = \Pr_{t_i\sim D_i}[t_i \le c_i]$ and $F_i(c_i \mid c_{-i}) = \Pr_{t_i \sim D_{i|c_{-i}}}[t_i \le c_i]$. Finally, when the buyer's type is $\boldsymbol{t}$, her valuation for a set of items S is denoted by $v(\boldsymbol{t}, S)$.

We investigate the performance of simple mechanisms for several well-studied valuation classes.

Definition 1 (Valuation Classes)**.** *We define several classes of valuations formally.*

- ***Constrained Additive:*** *interpret t_i as the value of item i, and $v(\boldsymbol{t}, S) = \max_{R\subseteq S, R\in\mathcal{I}} \sum_{i\in R} t_i$, where $\mathcal{I} \subseteq 2^{[m]}$ is a downward closed set system over the items specifying the feasible bundles. When $\mathcal{I} = 2^{[m]}$, the valuation is called **Additive**. When $\mathcal{I}$ contains all the singletons and the empty set, the valuation is called **unit-demand**.*

- ***XOS/Fractionally Subadditive:*** *interpret t_i as $\{t_i^{(k)}\}_{k\in[K]}$ that encodes all the possible values associated with item i, and $v(t, S) = \max_{k\in[K]} \sum_{i\in S} t_i^{(k)}$.*

It is well known that the class of XOS valuations contains all constrained additive valuations.

Mechanism A mechanism M is specified by an allocation rule and a payment rule. We use π to denote the allocation rule, and $\pi_i(\boldsymbol{t})$ is the probability that the buyer receives item i when she reports type $\boldsymbol{t}$. We also use $p(\boldsymbol{t})$ to denote the buyer's payment when she reports type $\boldsymbol{t}$. We assume the buyer has quasi-linear utility. We say a mechanism M is Incentive Compatible (IC) if the buyer cannot increase their expected utility by misreporting their type, and Individual Rational (IR) if the buyer has non-negative expected utility when they report their type truthfully to the mechanism.

Given D, valuation function $v(\cdot,\cdot)$, we use $\mathbf{Rev}(M, v, D)$ to denote the expected revenue of an IC and IR mechanism M. We slightly abuse notation to use $\mathbf{Rev}(D)$ to denote the optimal revenue achievable by any IC and IR mechanism under distribution D.

Throughout the paper, we use the following notations for the simple mechanisms we consider.

- SRev(v, D) denotes the optimal expected revenue achievable by any posted price mechanism, and we use SREV for short if there is no ambiguity. **- BRev**(v, D) denotes the optimal expected revenue achievable by selling a grand bundle and we use BREV for short if there is no ambiguity.

2.1 Markov Random Fields

Definition 2 (Markov Random Fields)**.** *A Markov Random Field (MRF) is defined by a hypergraph $G = (V, E)$. Associated with every vertex $v \in V$ is a random variable X_v taking values in some alphabet Σ_v, as well as a potential function $\psi_v : \Sigma_v \to \mathbb{R}$. Associated with every hyperedge $e \subseteq V$ is a potential function $\psi_e : \Sigma_e \to \mathbb{R}$. In terms of these potentials, we define a probability distribution π associating to each vector $\boldsymbol{c} \in \bigtimes_{v\in V} \Sigma_v$ probability $\pi(\boldsymbol{c})$ satisfying: $\pi(\boldsymbol{c}) \propto \prod_{v\in V} e^{\psi_v(c_v)} \prod_{e\in E} e^{\psi_e(\boldsymbol{c}_e)}$,where Σ_e denotes $\times_{v\in e}\Sigma_v$ and $\boldsymbol{c}_e$ denotes $\{c_v\}_{v\in e}$.*

We refer the interested readers to [35, 32] and the references therein for more details about MRFs. Throughout the paper, when we say the type distribution D is a MRF over a hypergraph $G = (V, E)$, if $V = [n]$, $t_i = x_i$, $T_i = \Sigma_i$, and there exists a collection of potential functions $\{\psi_i(\cdot)\}_{i\in[n]}$ and $\{\psi_e(\cdot)\}_{e\in E}$ so that the corresponding distribution $p(\cdot)$ equals to D. If there are only pairwise potentials, then G is a graph. We say that a random variable $\boldsymbol{t}$ is generated by a MRF, if $\boldsymbol{t}$ is sampled from a distribution that is represented by the MRF.

Next, we define two ways to measure the degree of dependence in a MRF.

Definition 3. *Let random variable $\boldsymbol{t}$ be generated by a Markov Random Field over a hypergraph $G = ([n], E)$, we define the* ***Markov influence*** *between item i and j to be: $\beta_{i,j}(\boldsymbol{t}) := \max_{\boldsymbol{x}\in\times_{\ell\in[n]}T_\ell} \left|\sum_{\substack{e\in E:\\ i,j\in e}} \psi_e(\boldsymbol{x}_e)\right|$. We further define the* ***inverse temperature*** *of the MRF as $\beta(\boldsymbol{t}) := \max_{i\in[n]} \sum_{j\neq i} \beta_{i,j}(\boldsymbol{t})$. We say random variable/type $\boldsymbol{t}$ is in the* ***high temperature regime*** *if $\beta(\boldsymbol{t}) < 1$.*

Definition 4. *Given a random variable/type $\boldsymbol{t}$ generated by a Markov Random Field over a hypergraph $G = ([n], E)$, we define the* ***weighted degree*** *of item i as: $d_i(\boldsymbol{t}) := \max_{\boldsymbol{x}\in\times_{i\in[n]}T_i} \left|\sum_{e\in E: i\in e} \psi_e(\boldsymbol{x}_e)\right|$, and the* ***maximum weighted degree*** *as $\Delta(\boldsymbol{t}) := \max_{i\in[n]} d_i(\boldsymbol{t})$.*

Remark 1. *Both $\beta(\boldsymbol{t})$ and $\Delta(\boldsymbol{t})$ capture the degree of dependence between the items. Note that $\Delta(\boldsymbol{t}) \leq \beta(\boldsymbol{t})$ for any MRF $\boldsymbol{t}$, and it is possible that $\beta(\boldsymbol{t}) = \Omega(d \cdot \Delta(\boldsymbol{t}))$, where d is the size of the largest hyperedge in G. When $\boldsymbol{t}$ is drawn from a product measure, both $\beta(\boldsymbol{t})$ and $\Delta(\boldsymbol{t})$ are 0. In general, restricting $\beta(\boldsymbol{t})$ and $\Delta(\boldsymbol{t})$ to be small ensures that the item values are only weakly dependent.*

To achieve our results, we need another important concept – the Glauber dynamics. In Section 5, we relate the approximation ratio achievable by simple mechanisms to the spectral gap of the Glauber dynamics of the MRF.

Definition 5 (Glauber Dynamics)**.** *Let $X_1, \ldots, X_n$ be an n-dimensional random vector drawn from distribution π. Let Ω be the support of π. The Glauber dynamics for π is a reversible Markov chain with state space Ω. The Glauber chain moves from state $x \in \Omega$ as follows: an index i is chosen uniformly at random from $[n]$, and a new state y is chosen so that (i) $y_j = x_j$ for all $j \neq i$; (ii) draw y_i from the conditional distribution $\pi \mid X_{-i} = x_{-i}$. It is not hard to verify that the Glauber dynamics is a reversible Markov chain with stationary distribution π.*

Remark 2. *When π is the distribution that can be represented by a MRF $G = (V, E)$, the Glauber dynamics has state space $\bigtimes_{v\in V} \Sigma_v$. The Glauber chain moves from state $x \in \bigtimes_{v\in V} \Sigma_v$ as follows:*

a vertex v is chosen uniformly at random from V, and a new state y is chosen so that (i) $y_u = x_u$ for all $u \neq v$; (ii) for any $c \in \Sigma_v$, $y_v = c$ w.p. $\frac{\exp(\psi_v(c))\Pi_{e:v\in e}\exp(\psi_e(c,x_{e/\{v\}}))}{\sum_{c'\in\Sigma_v}\exp(\psi_v(c'))\Pi_{e:v\in e}\exp(\psi_e(c',x_{e/\{v\}}))}$, in other words, sample y_v according to the distribution conditioned on $y_{-v} = x_{-v}$. Note that for a MRF, the Glauber dynamics is an irreiducible Markov chain, so π is its only stationary distribution. The Glauber dyanamics is a standard method for generating samples from a MRF, as it does not require computing the partition function, which is often a computationally intractable task.

3 Markov Random Fields: Basic Properties and Tools

We first present some basic properties of a MRF. Roughly speaking, we show that the conditional distribution can be approximated by the corresponding marginal distribution of D, and the approximation quality only depends $\Delta(\boldsymbol{t})$.

Lemma 1. *Let random variable $\boldsymbol{t}$ be generated by a MRF. Then for any $t_i \in T_i, t_{-i} \in T_{-i}$:*

$$\frac{\exp(\psi_i(t_i))}{\sum_{t'_i\in T_i}\exp(\psi_i(t'_i))}\exp(-2\Delta(\boldsymbol{t})) \le f_i(t_i \mid t_{-i}) \le \frac{\exp(\psi_i(t_i))}{\sum_{t'_i\in T_i}\exp(\psi_i(t'_i))}\exp(2\Delta(\boldsymbol{t}))$$

and

$$f_i(t_i)\cdot\exp(-4\Delta(\boldsymbol{t})) \le f_i(t_i \mid t_{-i}) \le f_i(t_i)\cdot\exp(4\Delta(\boldsymbol{t})).$$

Proof. Note that for any $t'_i \in T_i$, $\frac{f((t_i,t_{-i}))}{f((t'_i,t_{-i}))} = \frac{\exp(\psi_i(t_i))}{\exp(\psi_i(t'_i))}\cdot\frac{\exp\left(\sum_{e\in E, i\in e}\psi_e(t_i,t_{-i})_e\right)}{\exp\left(\sum_{e\in E, i\in e}\psi_e(t'_i,t_{-i})_e\right)}$.

Clearly,

$$\frac{\exp(\psi_i(t_i))}{\exp(\psi_i(t'_i))}\cdot\exp(-2\Delta(\boldsymbol{t})) \le \frac{f\left((t_i,t_{-i})\right)}{f\left((t'_i,t_{-i})\right)} \le \frac{\exp(\psi_i(t_i))}{\exp(\psi_i(t'_i))}\cdot\exp(2\Delta(\boldsymbol{t})).$$

Since $f_i(t_i \mid t_{-i}) = \frac{f((t_i,t_{-i}))}{\sum_{t'_i\in T_i} f\left((t'_i,t_{-i})\right)}$,

$$\frac{\exp(\psi_i(t_i))\cdot\exp(-2\Delta(\boldsymbol{t}))}{\sum_{t'_i\in T_i}\exp(\psi_i(t'_i))} \le f_i(t_i \mid t_{-i}) \le \frac{\exp(\psi_i(t_i))\exp(2\Delta(\boldsymbol{t}))}{\sum_{t'_i\in T_i}\exp(\psi_i(t'_i))}.$$

By Law of Total Probability,

$$f_i(t_i) = \sum_{t_{-i}\in T_{-i}} f_i(t_i \mid t_{-i})f_{-i}(t_{-i}) \in \left[\frac{\exp(\psi_i(t_i))\cdot\exp(-2\Delta(\boldsymbol{t}))}{\sum_{t'_i\in T_i}\exp(\psi_i(t'_i))}, \frac{\exp(\psi_i(t_i))\cdot\exp(2\Delta(\boldsymbol{t}))}{\sum_{t'_i\in T_i}\exp(\psi_i(t'_i))}\right].$$

□

Lemma 2. *Let random variable $\boldsymbol{t}$ be generated by a MRF. For any i and any set $\mathcal{E} \subseteq T_i$ and set $\mathcal{E}' \subseteq T_{-i}$:*

$$\exp(-4\Delta(\boldsymbol{t})) \le \frac{\Pr_{\boldsymbol{t}\sim D}\left[t_i \in \mathcal{E} \wedge t_{-i} \in \mathcal{E}'\right]}{\Pr_{t_i\sim D_i}[t_i \in \mathcal{E}]\Pr_{t_{-i}\sim D_{-i}}\left[t_{-i} \in \mathcal{E}'\right]} \le \exp(4\Delta(\boldsymbol{t})).$$

Proof. Note that $\Pr_{\boldsymbol{t}\sim D}\left[t_i \in \mathcal{E} \wedge t_{-i} \in \mathcal{E}'\right] = \sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i \mid t_{-i})\cdot f_{-i}(t_{-i})$ and $\Pr[t_i \in \mathcal{E}]\Pr\left[t_{-i} \in \mathcal{E}'\right] = \sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})$. Hence

$$\frac{\Pr_{\boldsymbol{t}\sim D}\left[t_i \in \mathcal{E} \wedge t_{-i} \in \mathcal{E}'\right]}{\Pr_{t_i\sim D_i}[t_i \in \mathcal{E}]\Pr_{t_{-i}\sim D_{-i}}\left[t_{-i} \in \mathcal{E}'\right]} = \frac{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i \mid t_{-i})\cdot f_{-i}(t_{-i})}{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})}$$

Using Lemma 1 we get that:

$$\frac{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i \mid t_{-i})\cdot f_{-i}(t_{-i})}{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})} \leq \exp(4\Delta)\frac{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})}{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})} = \exp(4\Delta)$$

and

$$\frac{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i \mid t_{-i})\cdot f_{-i}(t_{-i})}{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})} \geq \exp(-4\Delta)\frac{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})}{\sum_{\boldsymbol{t}\in\mathcal{E}\times\mathcal{E}'} f_i(t_i)\cdot f_{-i}(t_{-i})} = \exp(-4\Delta).$$

□

Prophet Inequality for MRF Equipped with Lemma 2, we provide a generalization of the Prophet inequality when the rewards in different stages are dependent and generated by a MRF. We can think of the prophet inequality problem, as finding a good policy for a gambler in a multi-round game. At the i-th round, the gambler is given the choice to accept a reward or to continue to the next round. The goal of the gambler is to find a policy that obtains high expected reward, given the distributions of the rewards at each round. Prophet inequalities have been obtained when the rewards between stages are independent [34, 42, 33] or can be expressed a a linear combination of some independent random variables [31].

Lemma 3. *Let $\boldsymbol{t} = (t_1, \ldots, t_n)$ be an n-dimensional random vector generated by a MRF. There are totally n rounds, and the reward of round i is $g_i(t_i)$, where g_i is an arbitrary function. The total reward of the prophet is $\mathbb{E}_{\boldsymbol{t}}\left[\max_{i\in[n]} g_i(t_i)\right]$. We denote by $\textsc{Reward}_{\boldsymbol{t}}\left[\{g_i\}_{i\in[n]}, \tau\right]$ the expected of reward of the following policy – accept any reward that is at least τ. The following inequality holds if we choose $\tau^* = \textsc{Median}_{\boldsymbol{t}}\left(\max_{i\in[n]} g_i(t_i)\right)$ (i.e., $\Pr[\max_{i\in[n]} g_i(t_i) \geq \tau^*] = 1/2$),*

$$\frac{\exp(-4\Delta(\boldsymbol{t}))}{2}\,\mathbb{E}_{\boldsymbol{t}}\left[\max_{i\in[n]} g_i(t_i)\right] \leq \textsc{Reward}_{\boldsymbol{t}}\left[\{g_i\}_{i\in[n]}, \tau^*\right].$$

Proof. The proof is similar to the case when all t_i are independent.

It is not hard to see that

$$\mathbb{E}_{\boldsymbol{t}}\left[\max_{i\in[n]} g_i(t_i)\right] \leq \tau^* + \sum_{i\in[n]} \mathbb{E}_{t_i\sim D_i}\left[(g_i(t_i)-\tau^*)^+\right].$$

We provide a lower bound on $\textsc{Reward}_{\boldsymbol{t}}\left[\{g_i\}_{i\in[n]}, \tau^*\right]$.

$$\textsc{Reward}_{\boldsymbol{t}}\left[\{g_i\}_{i\in[n]}, \tau^*\right] \geq \Pr_{\boldsymbol{t}}\left[\max_{i\in[n]} g_i(t_i) \geq \tau^*\right]\cdot\tau^* + \sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}}\left[(g_i(t_i)-\tau^*)^+ \cdot \mathbb{1}[\max_{j\neq i} g_j(t_j) \leq \tau^*]\right]$$

For every $i \in [n]$, we define the set $\mathcal{E}_i$ as $\{t_i \in T_i : g_i(t_i) > \tau^*\}$ and $\mathcal{E}'_i$ as $\{t_{-i} \in T_{-i} : \max_{j\neq i} g_j(t_j) \leq \tau^*\}$. Note that

$$\begin{aligned}\mathbb{E}_{\boldsymbol{t}}\left[(g_i(t_i)-\tau^*)^+ \cdot \mathbb{1}[\max_{j\neq i} g_j(t_j) < \tau^*]\right] &= \sum_{\boldsymbol{t}\in\mathcal{E}_i\times\mathcal{E}'_i} (g_i(t_i)-\tau^*)\cdot f(\boldsymbol{t}) \\ &\geq \exp(-4\Delta(\boldsymbol{t}))\sum_{\boldsymbol{t}\in\mathcal{E}_i\times\mathcal{E}'_i} (g_i(t_i)-\tau^*)\cdot f_i(t_i)f_{-i}(t_{-i}) \\ &= \exp(-4\Delta(\boldsymbol{t}))\,\mathbb{E}_{t_i\sim D_i}\left[(g_i(t_i)-\tau^*)^+\right]\Pr_{t_{-i}\sim D_{-i}}[t_{-i}\in\mathcal{E}'_i]\end{aligned}$$

The inequality is due to Lemma 2. Putting everything together, we know that

$$\textsc{Reward}_{\boldsymbol{t}}\left[\{g_i\}_{i\in[n]},\tau^*\right] \geq \frac{1}{2}\cdot\tau^* + \sum_{i\in[n]} \mathbb{E}_{t_i\sim D_i}\left[(g_i(t_i)-\tau^*)^+\right]\cdot\frac{\exp(-4\Delta(\boldsymbol{t}))}{2},$$

which is at least $\frac{\exp(-4\Delta(\boldsymbol{t}))}{2}$ of the upper bound we provide for $\mathbb{E}_{\boldsymbol{t}}\left[\max_{i\in[n]} g_i(t_i)\right]$. [7] □

4 Simple Mechanisms for a Unit-Demand or Additive Buyer under MRF

In this section, we first use the duality framework from [14, 17] to construct an upper bound of $\textsc{Rev}(D)$. Next, we prove that if the buyer has either unit-demand or additive valuation across the items, $\max\{\textsc{SRev},\textsc{BRev}\}$ is a $O(\exp(12\Delta(\boldsymbol{t})))$-approximation or a $O(\exp(4\Delta(\boldsymbol{t})))$-approximation of $\textsc{Rev}(D)$, respectively.

4.1 Benchmark of the Optimal Revenue for Constrained Additive Valuations

In this section, we use the duality framework from [14, 17] to construct an upper bound of $\textsc{Rev}(D)$. We describe a benchmark of the optimal revenue for all constrained additive valuations. Deriving a benchmark for XOS valuations requires some extra care, and we provide details of the derivation in Section 5.1 when we study XOS valuations. We first remind the readers the partition of type space used in [14, 17].

Definition 6 (Partition of the Type Space for Constrained Additive Valuations [14, 17])**.** *We partition the type space T into n regions, where $R_i = \{\boldsymbol{t}\in T : i$ is the smallest index in $\text{argmax}_{i'\in[n]} t_{i'}\}$. If $\boldsymbol{t}\in R_i$, we call item i the **favorite item** of type $\boldsymbol{t}$.*

To handle the dependence across the items, we introduce some new notations to specify the benchmark.

Definition 7 (Ironed Virtual Value)**.** *Let D be the type distribution. For any $\boldsymbol{t}\in R_i$, we use $\phi_i(t_i)$ to denote the **ironed Myerson's virtual value** for distribution D_i, $\phi_i(t_i \mid t_{-i})$ to denote the **ironed Myerson's virtual value** when we ironed $D_{i|t_{-i}}$ over interval $[\max_{j\neq i} t_j, \max \textsc{Supp}(D_{i|t_{-i}})]$.*

If $D_{i|t_{-i}}$ is a regular distribution and $t_i' = \text{argmin}\{\hat{t}\in\textsc{Supp}(D_{i|t_{-i}}) : \hat{t} > t_i\}$,

$$\phi_i(t_i \mid t_{-i}) = t_i - \frac{(t_i'-t_i)\cdot\Pr_{\hat{\boldsymbol{t}}\sim D}\left[\hat{t}_i > t_i \wedge \hat{t}_{-i} = t_{-i}\right]}{f(\boldsymbol{t})} = t_i - \frac{(t_i'-t_i)\cdot(1-F_i(t_i\mid t_{-i}))}{f_i(t_i\mid t_{-i})}.$$

Moreover, $\phi_i(t_i \mid t_{-i})$ always satisfies the following property:

$$\max_{p\geq \max_{j\neq i} t_j} p\cdot(1-F_i(p\mid t_{-i})) = \sum_{t_i:\ (t_i,t_{-i})\in R_i} f_i(t_i\mid t_{-i})\cdot\phi_i(t_i\mid t_{-i})^+,$$

where $x^+ = \max\{x, 0\}$.

Lemma 4 contains the benchmark we use. See Appendix A for more details about Lemma 4.

[7] When $\max_{i\in[n]} g_i(t_i)$ is a discrete distribution, we may not be able to pick a τ^* such that $\Pr[\max_{i\in[n]} g_i(t_i) \geq \tau^*] = 1/2$, but it is folklore that this can be resolved by carefully picking a tie-breaking rule. We do not include the details here.

Lemma 4 (Benchmark of Optimal Revenue for Constrained Additive Valuations)**.** *Given a distribution D over the type space T, and a mechanism $M = (\pi, p)$, if the buyer's valuation v is constrained additive, then we have the following benchmark:*

$$\begin{aligned}
\textsc{Rev}(M, v, D) \leq & \sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot \phi_i(t_i \mid t_{-i}) \cdot \mathbb{1}\left[\boldsymbol{t} \in R_i\right] \quad (\textsc{Single}) \\
& + \sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot t_i \cdot \mathbb{1}\left[\boldsymbol{t} \notin R_i\right] (\textsc{Non-Favorite}) \\
\leq & \sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot \phi_i(t_i \mid t_{-i}) \cdot \mathbb{1}\left[\boldsymbol{t} \in R_i\right] \quad (\textsc{Single}) \\
& + \sum_{i \in [n]} \sum_{t_i > r} f_i(t_i) \cdot t_i \cdot \Pr_{\boldsymbol{t}' \sim D}\left[\boldsymbol{t}' \notin R_i \mid t_i' = t_i\right] \quad (\textsc{Tail}) \\
& + \sum_{i \in [n]} \sum_{t_i \leq r} f_i(t_i) \cdot t_i \quad (\textsc{Core}),
\end{aligned}$$

where $r = \textsc{SRev}(v, D)$.

Single-Dimensional Copies Setting: In the analysis of unit-demand bidders with independent items [19, 14], the optimal revenue is upper bounded by the optimal revenue in the single-dimensional copies setting defined in [19]. We make use of the same technique in our analysis. There is a single item for sale, and we construct n agents, where agent i has value t_i for winning the item. Notice that this is a single-dimensional setting, as each agent's type is specified by a single number.

4.2 A Unit-Demand Buyer

In this section, we show that a simple posted price mechanism can extract $O(\exp(12\Delta(\boldsymbol{t})))$ fraction of the optimal revenue when the type distribution D is a MRF. We first use the revenue of the Ronen's lookahead auction [40] to upper bound the benchmark from Lemma 4. [8] Ronen's auction first identifies the highest bidder, and offers a take it or leave it price to the highest bidder to maximize the revenue conditioned on the other bidders' types. The proof follows from the definition of Ronen's lookahead auction and basic properties of MRF presented in Lemma 1 and 2. We postpone the proof to Appendix B.

Lemma 5. *Let the type distribution D be represented by a MRF, M be any IC and IR mechanism for a unit-demand buyer, and $\textsc{Ronen}^{\textsc{Copies}}$ be revenue of the Ronen's lookahead auction [40] in the* $\textsc{Copies}$ *settings with respect to D. The following inequalities hold:*

$$\begin{aligned}
\max\{\textsc{Single}, \textsc{Non-Favorite}\} &\leq \textsc{Ronen}^{\textsc{Copies}} \\
\textsc{Ronen}^{\textsc{Copies}} &\leq \exp(8\Delta(\boldsymbol{t}))\, \mathbb{E}_{\boldsymbol{t}}\left[\max_{i \in [n]} \phi_i(t_i)^+\right].
\end{aligned}$$

Equipped with Lemma 5, we can apply the prophet inequality for MRF to show that a posted-price mechanism can obtain expected revenue that is at least $\Omega\left(\frac{\textsc{Ronen}^{\textsc{Copies}}}{\exp(12\Delta(\boldsymbol{t}))}\right)$. We delay the proof to Appendix B.

[8] Ronen's lookahead auction considers the setting where the seller is selling a single item to a set of buyers, whose values for the item are correlated.

Theorem 1. *Let the type distribution D be represented by a MRF. If the buyer's valuation is unit-demand, then there exists a posted-price mechanism M with prices $\{p_i\}_{i\in[n]}$ that obtains expected revenue at least $\frac{\textsc{Rev}(D)}{8\exp(12\Delta(\boldsymbol{t}))}$.*

Is it possible to improve the dependence on Δ? In Theorem 8, we show that if we use the optimal revenue in the COPIES setting as a benchmark of the optimal revenue in the original setting, then the exponential dependence on $\Delta(\boldsymbol{t})$ is unavoidable.

4.3 An Additive Buyer

In this section, we show that $\max\{\textsc{SRev}, \textsc{BRev}\}$ is a $O(\exp(4\Delta(\boldsymbol{t})))$ approximation of the optimal revenue when the type distribution D is a MRF. We denote by r_i the revenue of Myerson's auction for selling item i only. We use $r = \sum_{i\in[n]} r_i$ to denote $\textsc{SRev}$, as the revenue collected from item i only depends on the marginal distribution D_i. We first upper bound the terms $\textsc{Single}$ and $\textsc{Tail}$ by $\exp(4\Delta(t)) \cdot \textsc{SRev}$. The proof follows from a combination of the standard analysis of the terms $\textsc{Single}$ and $\textsc{Tail}$ from [14, 17] with properties of MRFs (Lemma 2). We postpone the proof to Appendix C.

Lemma 6. *Let the type distribution D be a MRF and M be any IC and IR mechanism for an additive buyer. The following inequalities holds:* $\textsc{Single} \leq \exp(4\Delta(\boldsymbol{t})) \cdot \textsc{SRev}$ *and* $\textsc{Tail} \leq \exp(4\Delta(\boldsymbol{t})) \cdot \textsc{SRev}$.

Finally, we analyze the $\textsc{Core}$. We define new random variables $C_i = t_i \cdot \mathbb{1}[t_i \leq r]$. Let $C = \sum_{i=1}^n C_i$. Note that $\mathbb{E}[C] = \textsc{Core}$. We first provide an upper bound on $\textsc{Var}[C]$, and show that if we sell the grand bundle at an appropriate price, its revenue is close to the $\textsc{Core}$. Note that under the item-independence assumption, it is not hard to show that $\textsc{Var}[C]$ is upper bounded by $2r^2$ [4, 14]. However, this analysis does not extend to the case where the buyer type is generated by a MRF. We first obtain a new upper bound of $\textsc{Var}[C]$. As $C = \sum_{i=1}^n C_i$, we have $\textsc{Var}[C] = \sum_{i\in[n]} \textsc{Var}[C_i] + \sum_{i\neq j} \textsc{Cov}[C_i, C_j]$. We further bound $\sum_{i\in[n]} \textsc{Var}[C_i]$ by $2r^2$ using the standard analysis in [4, 14] and each covariance $\textsc{Cov}[C_i, C_j]$ using properties of MRF (Lemma 2). The proof is postponed to Appendix C.

Lemma 7. *Let the type distribution D be represented by a MRF. For any $i, j \in [n]$, $\textsc{Cov}[C_i, C_j] \leq (\exp(4\Delta(\boldsymbol{t})) - 1)\,\mathbb{E}[C_i]\,\mathbb{E}[C_j]$. Moreover, $\textsc{Var}[C] \leq 2r^2 + (\exp(4\Delta(\boldsymbol{t})) - 1)\,\mathbb{E}[C]^2$.*

In the item-independence case, the standard analysis [4, 14] applies Chebyshev's inequality to show that the seller can sell the grand bundle at price $\mathbb{E}[C] - 2r$ with probability at least $1/2$, which implies that $\textsc{Core}$ is $O(\textsc{BRev} + \textsc{SRev})$. As our upper bound on $\textsc{Var}[C]$ is a lot larger, Chebyshev's inequality only gives a vacuous bound on the sell probability. [9] To show that selling the grand bundling is a good approximation of the $\textsc{Core}$, we set the price of the grand bundle differently and use the Paley-Zygmund inequality to prove that either the sell probability is high or the $\textsc{Core}$ is within a constant factor of r. The proof of Theorem 2 can be found in Appendix C.

Theorem 2. *Let the type distribution D be represented by a MRF. If the buyer's valuation is additive, then*

$$\left(2\exp(4\Delta(\boldsymbol{t})) + \sqrt{2}\right) \cdot \textsc{SRev} + 8\left(\exp(4\Delta(t)) + 1\right) \cdot \textsc{BRev} \geq \textsc{Rev}(D).$$

[9] In particular, if we set the price to be $a \cdot \mathbb{E}[C] - \kappa \cdot r$ for any constant $a \in [0, 1]$ and κ, Chebyshev's inequality tells us that the probability that the buyer cannot afford the grand bundle is at most $\frac{\textsc{Var}[C]}{((1-a)\,\mathbb{E}[C]+\kappa\cdot r)^2}$. However, our upper bound of $\textsc{Var}[C]$ will be larger than $((1-a)\,\mathbb{E}[C] + \kappa \cdot r)^2$, if $\exp(4\Delta(\boldsymbol{t})) > 2$ and $\mathbb{E}[C]$ is much larger than r. In this case, $\frac{\textsc{Var}[C]}{((1-a)\,\mathbb{E}[C]+\kappa\cdot r)^2}$ is larger than 1 making the bound useless.

In the following theorem, we show that the approximation ratio must have polynomial dependence on $\Delta(\boldsymbol{t})$. Our proof is based on a modification of the hard instance by Hart and Nisan [30]. They construct a joint distribution over two items with support size m and show that the optimal revenue is at least $m^{1/7} \cdot \max\{\text{BRev}, \text{SRev}\}$. Unfortunately, their construction requires Δ to be infinite. We show how to modify their construction so that the new distribution has maximum weighted degree $\Delta = O(m)$, and the gap between the optimal revenue and $\max\{\text{BRev}, \text{SRev}\}$ remains to be $m^{1/7}$. The key is to show that under the new distribution, no type shows up too rarely, and the optimal revenue, SRev, and BRev remain roughly the same. The proof is postponed to Appendix F.

Theorem 3. *For any sufficiently large $m \in \mathbb{N}$, there exists a type distribution over two items represented by a MRF D such that (i) the maximum weighted degree Δ is at most $C \cdot m$, where C is an absolute constant; (ii) for an additive buyer whose type is sampled from D, there exists an absolute constant $C' > 0$ such that $\text{Rev}(D) \geq C' m^{1/7} \cdot \max\{\text{BRev}(D), \text{SRev}(D)\}$.*

5 Simple Mechanisms for a XOS Buyer

5.1 Duality Framework for XOS Valuations

The benchmark is obtained using essentially the same approach as in [17]. Suppose the buyer has a XOS valuation function $v(\boldsymbol{t}, S)$. We denote by $V_i(\boldsymbol{t}) = v(\boldsymbol{t}, \{i\})$. We abuse this notation and we also define for $t_i \in T_i$, $V_i(t_i) = v((\mathbf{0}, \ldots, t_i, \ldots, \mathbf{0}), \{i\})$, where $\mathbf{0}$ is the all 0 vector. We summarize the benchmark for a XOS buyer in the following Lemma. More details can be found in Appendix E.

Lemma 8. *Partition the type space T into n regions, where*

$$R_i := \{\boldsymbol{t} \in T : f(\boldsymbol{t}) > 0 \text{ and } i \text{ is the smallest index that belongs in } \operatorname{argmax}_{i \in [n]} V_i(\boldsymbol{t})\ \}$$

Let $r = \text{SRev}$ be the revenue of the optimal posted price mechanism that allows the buyer to purchase at most one item. Let $C(\boldsymbol{t}) := \{i : V_i(\boldsymbol{t}) < 2r\}$. For any IC and IR Mechanism M, we can bound its revenue by:

$$\begin{aligned} \text{Rev}(M, v, D) \leq & 2 \sum_{\boldsymbol{t} \in T} f(\boldsymbol{t}) \sum_{i \in [n]} \pi_i(\boldsymbol{t}) \phi(V_i(t_i) \mid \boldsymbol{t}_{-i}) \mathbb{1}[t \in R_i] \quad (\text{Single}) \\ & + 4 \sum_{i \in [n]} \sum_{\substack{t_i \in T_i \\ V_i(t_i) \geq 2r}} f(t_i) \cdot V_i(t_i) \Pr_{\boldsymbol{t}' \sim D} \left[\boldsymbol{t}' \notin R_i \mid t'_i = t_i\right] (\text{Tail}) \\ & + 4 \sum_{\boldsymbol{t} \in T} f(\boldsymbol{t}) v(\boldsymbol{t}, C(\boldsymbol{t})) \quad (\text{Core}) \end{aligned}$$

5.2 Approximating the Benchmark of a XOS Buyer

In this section, we show how to approximate the optimal revenue of a buyer with a XOS valuation. We first upper bound the term Single and Tail. The analysis of both terms follows from the combination of the analysis in [17] and Lemma 2.

Lemma 9. *Let the type distribution D be represented by a MRF. If M is an IC and IR mechanism for a buyer with a XOS valuation, then the following inequalities hold*

$$\text{Single} \leq 4 \exp(12\Delta(\boldsymbol{t})) \cdot \text{SRev}$$

and

$$\text{TAIL} \le \exp(8\Delta(\boldsymbol{t})) \cdot \text{SREV},$$

where SREV *is the revenue of the optimal posted price auction, in which the buyer is allowed to purchase at most one item.*

5.2.1 Bounding the Core using the Poincaré Inequality

In this section, we show how to bound the CORE for a XOS buyer. The CORE is the expectation of the random variable $v(\boldsymbol{t}, C(\boldsymbol{t}))$. To show that bundling can achieve a good approximation of the CORE, we need to upper bound the variance of $v(\boldsymbol{t}, C(\boldsymbol{t}))$. This is the main task of this section. As $v(\boldsymbol{t}, \cdot)$ is not additive across the items, our method for the additive valuation (see Lemma 7) no longer applies. We provide a new approach that is based on the *Poincaré Inequality* and the *self-boundingness of XOS functions.* We first state the *Poincaré Inequality.*

Lemma 10 (The Poincaré Inequality (adapted from Lemma 13.12 of [36]))**.** *Let P be a reversible transition matrix on state space Ω with stationary distribution π. For any function $g : \Omega \to \mathbb{R}$, let*

$$\mathcal{E}(g) := \frac{1}{2} \sum_{x,y\in\Omega} [g(x) - g(y)]^2 \pi(x) P(x, y).$$

If $\text{VAR}_{x\sim\pi}[g(x)] > 0$*, then*

$$\frac{\mathcal{E}(g)}{\text{VAR}_{x\sim\pi}[g(x)]} \ge \gamma,$$

where γ is the spectral gap of P. [10] *Moreover, there exists a function $g^* : \Omega \to \mathbb{R}$, such that*

$$\frac{\mathcal{E}(g^*)}{\text{VAR}_{x\sim\pi}[g^*(x)]} = \gamma.$$

Next, we apply Lemma 10 to the Glauber dynamics of the MRF that generates the buyer's type.

Lemma 11. *Let D be the joint distribution of random variables $\boldsymbol{t} = (t_1, \ldots, t_n)$ and P be the transition matrix of the Glauber dynamics for D. For any function $g : T \to \mathbb{R}$, we have*

$$n\gamma \cdot \text{VAR}_{\boldsymbol{t}\sim D}[g(\boldsymbol{t})] \le \sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}\sim D} \left[\left(g(t_i, \boldsymbol{t}_{-i}) - \mathbb{E}_{t'_i\sim D_{i|\boldsymbol{t}_{-i}}}[g(t'_i, \boldsymbol{t}_{-i})] \right)^2 \right],$$

where γ is the spectral gap of P. Moreover, there exists a function $g^* : T \to \mathbb{R}$, such that the inequality is tight.

Remark 3. *Lemma 11 is a generalization of the well-known Efron-Stein inequality to dependent random variables. Indeed, when D is a product measure, γ is at least $1/n$ and we recover the Efron-Stein inequality. As we demonstrate in Section 6, γ is at least $\Omega(1/n)$ under many well-studied conditions of weak dependence, such as the Dobrushin uniqueness condition.*

[10] It is well-known that the largest eigenvalue of P is 1, and the spectral gap of P is the difference between P's largest and second largest eigenvalues.

Proof of Lemma 11: According to the definition of the Glauber dynamics, P is a reversible transition matrix on state space T with stationary distribution D. Lemma 10 states that

$$\gamma \cdot \textsc{Var}_{\boldsymbol{t}\sim D}[g(x)] \leq \frac{1}{2} \sum_{\boldsymbol{t},\boldsymbol{t}' \in T} [g(\boldsymbol{t}) - g(\boldsymbol{t}')]^2 \cdot f(\boldsymbol{t}) \cdot P(\boldsymbol{t},\boldsymbol{t}'). \tag{1}$$

By the definition of the Glauber dynamics, the RHS of Inequality (1) is equivalent to

$$\begin{aligned}
&\frac{1}{2}\,\mathbb{E}_{\boldsymbol{t}\sim D}\left[\frac{1}{n}\sum_{i\in[n]} \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}\left[g(\boldsymbol{t}) - g(t_i', \boldsymbol{t}_{-i})\right]^2\right] \\
=&\frac{1}{n}\sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}_{-i}\sim D_{-i}}\left[\mathbb{E}_{t_i,t_i'\sim D_{i|\boldsymbol{t}_{-i}}}\left[\frac{1}{2}\left(g(t_i,\boldsymbol{t}_{-i}) - g(t_i',\boldsymbol{t}_{-i})\right)^2\right]\right] \\
=&\frac{1}{n}\sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}_{-i}\sim D_{-i}}\left[\mathbb{E}_{t_i\sim D_{i|\boldsymbol{t}_{-i}}}\left[\left(g(t_i,\boldsymbol{t}_{-i}) - \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}[g(t_i',\boldsymbol{t}_{-i})]\right)^2\right]\right] \\
=&\frac{1}{n}\sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}\sim D}\left[\left(g(t_i,\boldsymbol{t}_{-i}) - \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}[g(t_i',\boldsymbol{t}_{-i})]\right)^2\right].
\end{aligned}$$

The second equality is because t_i and t_i' are two i.i.d. samples from $D_{i|\boldsymbol{t}_{-i}}$.

Hence,

$$n\gamma \cdot \textsc{Var}_{\boldsymbol{t}\sim D}[g(\boldsymbol{t})] \leq \sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}\sim D}\left[\left(g(t_i,\boldsymbol{t}_{-i}) - \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}[g(t_i',\boldsymbol{t}_{-i})]\right)^2\right].$$

Note that if we choose $g(\cdot)$ to be the function $g^*(\cdot)$ in Lemma 10, Inequality (1) becomes an equality.

□

Recall that to bound the Core, we need to upper bound the variance of the random variable $v(\boldsymbol{t}, C(\boldsymbol{t}))$. By choosing $g(\boldsymbol{t})$ to be $v(\boldsymbol{t}, C(\boldsymbol{t}))$ and applying Lemma 11, we can instead upper bound the RHS of the inequality in Lemma 11. A priori, it is not clear that the RHS would be easier to bound. In the following sequence of Lemmas, we show that the RHS is indeed more amenable to analysis. We first argue that the function $v(\boldsymbol{t}, C(\boldsymbol{t}))$ has a key property known as *self-boundingness*, using which we then upper bound the RHS by $O(\textsc{SRev} \cdot \textsc{Core})$ and show that SRev and BRev can approximate the Core.

Definition 8 (Self-Bounding Functions [7])**.** *Let $\mathbb{S}$ be an arbitrary set and A be a subset of $\mathbb{S}^n$. We say that a function $g(\boldsymbol{t}) : A \to \mathbb{R}$ is* C-self-bounding *with some constant $C \in \mathbb{R}_+$ if there exists a collection of functions $g_i : A_{-i} \to \mathbb{R}$ for each $i \in [n]$ with $A_{-i} := \{\boldsymbol{t}_{-i} : \exists t_i,\ (t_i, \boldsymbol{t}_{-i}) \in A\}$, such that for each $\boldsymbol{t} \in A$ the followings hold:*

- $0 \leq g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i}) \leq C$ *for all* $i \in [n]$.
- $\sum_{i\in[n]} (g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i})) \leq g(\boldsymbol{t})$.

We next argue that for a self-bounding function, the RHS of the inequality in Lemma 11 is upper bounded by its mean.

Lemma 12. *Let D be the joint distribution of random variables $\boldsymbol{t} = (t_1, \ldots, t_n)$. If $g(\cdot)$ is a C-self-bounding function, then*

$$\sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}\sim D}\left[\left(g(\boldsymbol{t}) - \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}[g(t_i',\boldsymbol{t}_{-i})]\right)^2\right] \leq C\,\mathbb{E}_{\boldsymbol{t}\sim D}\left[g(\boldsymbol{t})\right].$$

Proof. Recall the following property of the variance: For any real-value random variable X, $\textsc{Var}[X] = \min_{a \in R} \mathbb{E}[(X-a)^2]$. In other words, $\textsc{Var}[X] \leq \mathbb{E}[(X-a)^2]$ for any a. Therefore, for any $\boldsymbol{t}_{-i}$,

$$\mathbb{E}_{t_i \sim D_{i|\boldsymbol{t}_{-i}}}\left[\left(g(\boldsymbol{t}) - \mathbb{E}_{t'_i \sim D_{i|\boldsymbol{t}_{-i}}}[g(t'_i, \boldsymbol{t}_{-i})]\right)^2\right] = \textsc{Var}[g(\boldsymbol{t}) \mid \boldsymbol{t}_{-i}] \leq \mathbb{E}_{t_i \sim D_{i|\boldsymbol{t}_{-i}}}\left[(g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i}))^2\right].$$

Using this relaxation, we proceed to prove the claim.

$$\begin{aligned}
&\sum_{i \in [n]} \mathbb{E}_{\boldsymbol{t} \sim D}\left[\left(g(\boldsymbol{t}) - \mathbb{E}_{t'_i \sim D_{i|\boldsymbol{t}_{-i}}}[g(t'_i, \boldsymbol{t}_{-i})]\right)^2\right] \\
\leq &\sum_{i \in [n]} \mathbb{E}_{\boldsymbol{t} \sim D}\left[(g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i}))^2\right] \\
\leq & C \sum_{i \in [n]} \mathbb{E}_{\boldsymbol{t} \sim D}\left[(g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i}))\right] \\
\leq & C\, \mathbb{E}_{\boldsymbol{t} \sim D}\left[g(\boldsymbol{t})\right]
\end{aligned}$$

The first inequality follows from the relaxation. The second and last inequality follow from the first and second property of a self-bounding function respectively. □

Combining Lemma 11 and 12, we have the following Lemma.

Lemma 13. *Let D be the joint distribution of random variables $\boldsymbol{t} = (t_1, \ldots, t_n)$ and P be the transition matrix of the Glauber dynamics for D. For any C-self-bounding function $g : T \to \mathbb{R}$, we have*

$$\frac{n\gamma}{C} \cdot \textsc{Var}_{\boldsymbol{t} \sim D}[g(\boldsymbol{t})] \leq \mathbb{E}_{\boldsymbol{t} \sim D}\left[g(\boldsymbol{t})\right],$$

where γ is the spectral gap of P.

Definition 8 may seem obscure at first, but many natural functions are indeed self-bounding. For example, if A is $[0,1]^n$ and $g(\cdot)$ is the additive function, then $g(\cdot)$ is 1-self-bounding. We show that the function $g(\boldsymbol{t}) := v(\boldsymbol{t}, C(\boldsymbol{t}))$ is $2\textsc{SRev}$-self-bounding and its variance is no more than $\frac{2\textsc{SRev} \cdot \textsc{Core}}{n\gamma}$. Here, we first specialize our analysis to MRFs. The main difference is that the Glauber dynamics for a MRF is irreducible, so the spectral gap is strictly positive (Lemma 12.1 of [36]). The proof is postponed to Appendix D.

Lemma 14. *Let $C(\boldsymbol{t}) := \{j : V_j(\boldsymbol{t}) < 2\textsc{SRev}\}$. The function $g(\boldsymbol{t}) := v(\boldsymbol{t}, C(\boldsymbol{t}))$ is $2\textsc{SRev}$-self-bounding and $\textsc{Var}_{\boldsymbol{t} \sim D}[g(\boldsymbol{t})] \leq \frac{2\textsc{SRev} \cdot \textsc{Core}}{n\gamma}$, where $\gamma > 0$ is the spectral gap of the transition matrix of the Glauber dynamics of the MRF that generates the buyer's type.*

Now, we show how to approximate $\textsc{Core}$ using $\textsc{SRev}$ and $\textsc{BRev}$.

Lemma 15. *Let the buyer's type distribution D be represented by a MRF, P be the transition matrix of the Glauber dynamics of the MRF, and $\gamma > 0$ be the spectral gap of P. We have*

$$\textsc{Core} \leq \max\left(\frac{4\textsc{SRev}}{\sqrt{n\gamma}}, \left(7 + \frac{4}{\sqrt{n\gamma}}\right)\textsc{BRev}\right).$$

Proof. According to Lemma 14, $v(\boldsymbol{t}, C(\boldsymbol{t}))$ is a 2SRev-self-bounding function and $\textsc{Var}[v(\boldsymbol{t}, C(\boldsymbol{t}))] \leq \frac{2\text{SRev}\cdot\textsc{Core}}{n\gamma}$. If $\textsc{Core} \leq \frac{4\text{SRev}}{\sqrt{n\gamma}}$, then the statement holds. If $\textsc{Core} > \frac{4\text{SRev}}{\sqrt{n\gamma}}$, then $\textsc{Var}[v(\boldsymbol{t}, C(\boldsymbol{t}))] \leq \frac{2\text{SRev}\cdot\textsc{Core}}{n\gamma} < \frac{(\textsc{Core})^2}{2\sqrt{n\gamma}} = \frac{\mathbb{E}[v(\boldsymbol{t},C(\boldsymbol{t}))]^2}{2\sqrt{n\gamma}}$. By Paley-Zygmund inequality we have that:

$$\Pr\left[v(\boldsymbol{t}, C(\boldsymbol{t})) \geq \frac{\textsc{Core}}{3}\right] \geq \frac{4}{9}\frac{1}{1+\frac{\textsc{Var}[v(\boldsymbol{t},C(\boldsymbol{t}))]}{\mathbb{E}[v(\boldsymbol{t},C(\boldsymbol{t}))]^2}} \geq \frac{4}{9}\frac{1}{1+\frac{1}{2\sqrt{n\gamma}}}.$$

Therefore we have that: $\Pr\left[v(\boldsymbol{t}, C(\boldsymbol{t})) \geq \frac{\textsc{Core}}{3}\right] \cdot \frac{\textsc{Core}}{3} \leq \text{BRev}$, which implies that the statement. □

Finally, we combine our analysis of Single, Tail, and Core to obtain the approximation guarantee for a XOS buyer.

Theorem 4. *Let the buyer have a XOS valuation and her type distribution D be represented by a MRF. We use γ to denote the spectral gap of matrix P – the transition matrix of the Glauber dynamics of the MRF. Then* $\text{Rev}(D) \leq 12\exp(12\Delta(\boldsymbol{t}))\cdot\text{SRev} + \left(28 + \frac{16}{\sqrt{n\gamma}}\right)\max\{\text{SRev}, \text{BRev}\}$.

Proof of Theorem 4: The statement follows from the combination of Lemma 8 , 9, and 15. □

6 Connection to other Weak Dependence Conditions

A common way to measure the degree of dependence of a high-dimensional distribution is by considering its *Dobrushin Interdependence Matrix*. In this section, we show that for several natural sufficient conditions that guarantee weak dependence in the distribution, the spectral gap γ of the Glauber dynamics transition matrix is $\Omega(1/n)$. We begin by defining the *Dobrushin interdependence matrix.*

Definition 9 (d-Dobrushin Interdependence Matrix [44])**.** *Let $(\mathbb{E}, d)$ be a metrical, complete and separable space. For two distributions μ and ν supported on $\mathbb{E}$, their L^1-Wasserstein distance is defined as: $W_{1,d}(\mu,\nu) = \inf_{\pi\in\Pi}\int\int_{\mathbb{E}\times\mathbb{E}} d(x,y)\pi(dx,dy)$, where Π is the set of valid coupling such that its marginal distributions are μ and ν.*

*Let $X = (x_1,\ldots,x_n)$ be a n-dimensional random vector supported on $\mathbb{E}^n$ and $\mu_i(\cdot \mid x_{-i})$ be the conditional distribution of x_i knowing x_{-i}. Define the d-**Dobrushin Interdependence Matrix** $A = (\alpha_{i,j})_{i,j\in[n]}$ by*

$$\alpha_{i,j} := \sup_{\substack{x_{-i-j}=y_{-i-j}\\ x_j\neq y_j}} \frac{W_{1,d}(\mu_i(\cdot \mid x_{-i}), \mu_i(\cdot \mid y_{-i}))}{d(x_j, y_j)} \text{ for all } i \neq j,$$

and $\alpha_{i,i} = 0$ for all $i \in [n]$.

Remark 4. *$\alpha_{i,j}$ captures how strong the value of x_j affects the conditional distribution of x_i when all other coordinates are fixed. Higher $\alpha_{i,j}$ value implies stronger dependence between x_i and x_j. When all the coordinates of X are independent, A is the all zero matrix.*

Dobrushin uniqueness condition: If we choose $d(x,y)$ to be the trivial metric $\mathbb{1}_{x\neq y}$, then $W_{1,d}(\cdot,\cdot)$ is exactly the total variation distance. The influence matrix mentioned in Section 1 is exactly the Dobrushin interdependence matrix with respect to the trivial metric. To remind the audience, the **Dobrushin Coefficient** is defined as $\alpha(\boldsymbol{t}) := ||A||_\infty = \max_{i\in[n]}\sum_{j\neq i}\alpha_{i,j}$ when A is the influence matrix. If $\alpha(\boldsymbol{t}) < 1$, we say $\boldsymbol{t}$ satisfies the **Dobrushin uniqueness condition**. As $||A||_\infty$ is at least as large as A's spectral radius $\rho_d(\boldsymbol{t})$, [11] a weaker condition than the Dobrushin uniqueness condition is that the spectral radius $\rho_d(\boldsymbol{t})$ is strictly less than 1.

We argue that even the weaker condition that $\rho_d(\boldsymbol{t}) < 1$ implies that the spectral gap of the transition matrix of the Glauber dynamics $\gamma = \Omega(1/n)$.

Lemma 16. *Let $d(\cdot,\cdot)$ be any metric, for any n-dimensional random vector $\boldsymbol{t}$, $n\gamma \geq 1-\rho_d(\boldsymbol{t})$, where γ is the spectral gap of the transition matrix of the Glauber dynamics for $\boldsymbol{t}$.*

Proof. By Lemma 11, there exists a function $g^*: \Omega \to \mathbb{R}$, such that

$$\frac{\sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}\sim D}\left[\left(g^*(t_i,\boldsymbol{t}_{-i}) - \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}[g^*(t_i',\boldsymbol{t}_{-i})]\right)^2\right]}{\textsc{Var}_{\boldsymbol{t}\sim D}[g^*(\boldsymbol{t})]} = n\gamma.$$

The following result by Wu [44] provides a generalization of the Efron-Stein inequality for *weakly dependent random variables.*

Lemma 17 (Poincaré Inequality for Weakly Dependent Random Variables - Theorem 2.1 in [44])**.** *Let $\boldsymbol{t} = (t_1,\dots,t_n)$ be an n-dimensional random vector drawn from distribution D that is supported on $\mathbb{E}^n$. For any metric $d(\cdot,\cdot)$ on $\mathbb{E}$, let A be the d-Dobrushin interdependence matrix for $\boldsymbol{t}$. Let $\rho_d(\boldsymbol{t})$ be the spectral radius of A. If $\mathbb{E}_{\boldsymbol{t}\sim D}\left[\sum_{i\in[n]} d(t_i,y_i)^2\right] < +\infty$ for some fixed $y \in \mathbb{E}^n$ and $\rho_d(\boldsymbol{t}) < 1$, then for any square integrable function $g(\cdot)$ w.r.t. distribution D, the following holds:*

$$(1-\rho_d(\boldsymbol{t}))\textsc{Var}_{\boldsymbol{t}\sim D}[g(\boldsymbol{t})] \leq \sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}\sim D}\left[\left(g(t_i,\boldsymbol{t}_{-i}) - \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}[g(t_i,\boldsymbol{t}_{-i})]\right)^2\right].$$

If we choose $g(\cdot)$ to be $g^*(\cdot)$ in Lemma 17, we have that :

$$1-\rho_d(\boldsymbol{t}) \leq \frac{\sum_{i\in[n]} \mathbb{E}_{\boldsymbol{t}\sim D}\left[\left(g^*(t_i,\boldsymbol{t}_{-i}) - \mathbb{E}_{t_i'\sim D_{i|\boldsymbol{t}_{-i}}}[g^*(t_i,\boldsymbol{t}_{-i})]\right)^2\right]}{\textsc{Var}_{\boldsymbol{t}\sim D}[g^*(\boldsymbol{t})]}$$

Combining the two inequalities conclude the proof □

Combining Lemma 16 with Theorem 4, we immediately have the following Theorem. [12]

Theorem 5. *Let the buyer have a XOS valuation, her type distribution D be represented by a MRF, and $\rho_d(\boldsymbol{t})$ be the spectral radius of the d-Dobrushin interdependence matrix of $\boldsymbol{t}$ under some metric $d(\cdot,\cdot)$. If $\rho_d(\boldsymbol{t}) < 1$, then* $\textsc{Rev}(D) \leq 12\exp(12\Delta(\boldsymbol{t}))\cdot \textsc{SRev} + \left(28 + \frac{16}{\sqrt{1-\rho_d(\boldsymbol{t})}}\right)\max\{\textsc{SRev}, \textsc{BRev}\}$.

Proof of Theorem 5: The statement follows from the combination of Lemma 16 and Theorem 4. □

[11] $\rho_d(\boldsymbol{t})$ is the dominant eigenvalue of A by the Perron-Frobenius Theorem.

[12] A major benefit of using ρ_d or the Dobrushin coefficient rather than γ is that these parameters are easier to estimate than γ given the joint distribution.

High Temperature MRFs. Using Theorem 5, we show that when the MRF is in the *high temperature regime*, i.e., $\beta(\boldsymbol{t}) < 1$ (see Definition 3), $\max\{\text{SRev}, \text{BRev}\}$ is a constant factor approximation to the optimal revenue. By the definition of $\beta(\boldsymbol{t})$, it clear that $\Delta(\boldsymbol{t}) \le \beta(\boldsymbol{t})$. Next, we show that $\beta(\boldsymbol{t})$ is also an upper bound of $\rho_d(\boldsymbol{t})$ for the trivial metric $d(x, y) = \mathbb{1}_{x \neq y}$.

Lemma 18. *Let $d(\cdot,\cdot)$ be the trivial metric $d(x, y) = \mathbb{1}_{x\neq y}$. For any MRF $\boldsymbol{t}$, $\rho_d(\boldsymbol{t}) \le \alpha(\boldsymbol{t}) \le \beta(\boldsymbol{t})$. Moreover, $\alpha_{i,j} \le \beta_{i,j}(\boldsymbol{t})$ for all $i, j \in [n]$.*

Proof. $\rho_d(\boldsymbol{t}) \le \alpha(\boldsymbol{t})$ follows the elementary fact that the spectral radius is upper bounded by the infinity norm. To prove $\alpha(\boldsymbol{t}) \le \beta(\boldsymbol{t})$, we first need the following definition and lemma.

Definition 10. *[22] Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a random variable over $\times_{i\in[n]}\Sigma_i$, and $P_{\boldsymbol{x}}$ denote its probability distribution. Assume $P_{\boldsymbol{x}} > 0$ on all $\times_{i\in[n]}\Sigma_i$. For any $i \neq j \in [n]$, define the log influence between i and j as*

$$I^{\log}_{i,j}(\boldsymbol{x}) = \frac{1}{4} \max_{\substack{x_{-i-j}\in\Sigma_{-i-j}\\ x_i,x'_i\in\Sigma_i\\ x_j,x'_j\in\Sigma_j}} \log \frac{P_{\boldsymbol{x}}(x_i x_j x_{-i-j}) P_{\boldsymbol{x}}(x'_i x'_j x_{-i-j})}{P_{\boldsymbol{x}}(x'_i x_j x_{-i-j}) P_{\boldsymbol{x}}(x_i x'_j x_{-i-j})}.$$

Lemma 19 (Adapted from Lemma 5.2 of [23])**.** *Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a random variable , for any $i, j \in [n]$, $\alpha_{i,j} \le I^{\log}_{i,j}(\boldsymbol{x})$.*

We only need to prove that $I^{\log}_{i,j}(\boldsymbol{t})$ is no more than $\beta_{i,j}(\boldsymbol{t})$. Since random variable $\boldsymbol{t}$ is generated by a MRF,

$$I^{\log}_{i,j}(\boldsymbol{t}) = \frac{1}{4} \max_{\substack{t_{-i-j}\in\Sigma_{-i-j}\\ t_i,t'_i\in\Sigma_i\\ t_j,t'_j\in\Sigma_j}} \sum_{\substack{e\in E:\\ i,j\in e}} \psi_e\left((t_i, t_j, t_{-i-j})_e\right) + \sum_{\substack{e\in E:\\ i,j\in e}} \psi_e\left((t'_i, t'_j, t_{-i-j})_e\right)$$
$$- \sum_{\substack{e\in E:\\ i,j\in e}} \psi_e\left((t'_i, t_j, t_{-i-j})_e\right) - \sum_{\substack{e\in E:\\ i,j\in e}} \psi_e\left((t_i, t'_j, t_{-i-j})_e\right),$$

which is clearly no greater than $\beta_{i,j}(\boldsymbol{t})$.

Since for any $i, j \in [n]$ $\alpha_{i,j} \le \beta_{i,j}(\boldsymbol{t})$, we have $\alpha(\boldsymbol{t}) \le \beta(\boldsymbol{t})$. □

Theorem 6. *Let the buyer's type distribution D be represented by a MRF. If the buyer's a XOS valuation and her type $\boldsymbol{t}$ is in the high temperature regime, i.e., $\beta(\boldsymbol{t}) < 1$,*

$$\text{Rev}(D) \le 12\exp(12\beta(\boldsymbol{t}))\cdot\text{SRev} + \left(28 + \frac{16}{\sqrt{1-\beta(\boldsymbol{t})}}\right)\max\{\text{SRev}, \text{BRev}\} = O\left(\frac{\max\{\text{SRev}, \text{BRev}\}}{\sqrt{1-\beta(\boldsymbol{t})}}\right).$$

Proof of Theorem 6: The statement follows from Theorem 5 and Lemma 18. □

7 Impossibility Results

In this section we present some of our impossibility results. In Section 7.1, we show that the Dobrushin Uniqueness condition alone is insufficient to guarantee any multiplicative approximation of the optimal revenue using SRev and BRev. In Section 7.2 we construct a MRF such that the optimal revenue in the COPIES setting is $\exp(\Delta)$ times larger than $\max\{\text{SRev}, \text{BRev}\}$.

7.1 Inapproximability under Only the Dobrushin Uniqueness Condition

Readers may wonder whether it is possible to prove an approximation ratio that only relies on either the spectral radius $\rho_d(\boldsymbol{t})$, the Dobrushin coefficient $\alpha(\boldsymbol{t})$, or the spectral gap of the Glauber dynamics γ, but independent of the maximum weighted degree $\Delta(\boldsymbol{t})$. We show that this is impossible. Indeed, we prove that for any $\alpha < 1$, and any ratio N, there exists a MRF with $\rho_d(\boldsymbol{t}) \le \alpha(\boldsymbol{t}) \le \alpha$ such that the ratio between the optimal revenue and $\max\{\text{BRev}(D), \text{SRev}(D)\}$ is at least $\frac{\alpha}{2} \cdot N$. Our result is based on a modification of the Hart-Nisan construction [30].

Theorem 7. *For any positive real number N and any choice of $0 < \alpha < 1$, there exists a type distribution D over 2 items generated by a MRF with Dobrushin coefficient $\alpha(\boldsymbol{t}) = \alpha$ and finite inverse temperature, such that for an additive buyer whose type is sampled from D, $\frac{\text{Rev}(D)}{\max\{\text{BRev}(D),\text{SRev}(D)\}} > \frac{\alpha}{2} \cdot N$.*

First we present the main building block of our construction.

Lemma 20. *Let D' be a correlated valuation distribution over 2 items with Dobrushin coefficient α. Let D be a product distribution that has the same marginal distributions as D'. Then for any $0 \le \alpha' \le 1$, we consider the distribution $D'' := \alpha' \cdot D' + (1 - \alpha') \cdot D$, that is, if we want to sample from D'', we can take a sample from D' with probability α' and take a sample from D with probability $1-\alpha'$. Distribution D'' can be modeled as a MRF with finite inverse temperature such that $\Delta = \beta(\boldsymbol{t}) \le |\log((1-\alpha)p^2)|$, where $p = \inf_{\boldsymbol{t}\in\text{Supp}(D')} \Pr_{\boldsymbol{t}'\sim D'}[\boldsymbol{t}' = \boldsymbol{t}]$ and D'' has Dobrushin coefficient $\alpha' \cdot \alpha$. Furthermore, D'' has the same marginal distribution as D and $\text{Rev}(D'') \ge \alpha'\text{Rev}(D')$.*

Proof. Assume that $D = D_1 \times D_2$, where D_1 and D_2 are the marginals of D'. Let $T_i = \text{Supp}(D_i)$and $T = T_1 \times T_2$. Let $A^{D'} = \{\alpha^{D'}_{i,j}\}_{i,j\in[2]}$ be the influence matrix of D' and $A^{D''} = \{\alpha^{D''}_{i,j}\}_{i,j\in[2]}$ be the influence matrix of D''. Note that the diagonal entries of $A^{D'}$ and of $A^{D''}$ are zero. We have that:

$$\alpha^{D''}_{i,j} = \max_{t_j,t'_j\in T_j} d_{TV}\left(D''_{i|t_j}, D''_{i|t'_j}\right) = \alpha' \max_{t_j,t'_j\in T_j} d_{TV}\left(D'_{i|t_j}, D'_{i|t'_j}\right) = \alpha'\alpha^{D'}_{i,j}$$

In the statement of the lemma, we assumed that $\alpha = ||A^{D'}||_\infty = \max(\alpha^{D'}_{1,2}, \alpha^{D'}_{2,1})$. We can easily infer the following: $||A^{D''}||_\infty = \max(\alpha^{D''}_{1,2}, \alpha^{D''}_{2,1}) = \alpha' \max(\alpha^{D'}_{1,2}, \alpha^{D'}_{2,1}) = \alpha' \cdot \alpha$. This concludes the fact that D'' has $\alpha(\boldsymbol{t}) = \alpha' \cdot \alpha$.

Now we prove that D'' can be modeled as a MRF with finite inverse temperature. We consider a MRF with potential functions $\psi_1(t_1) = 1, \forall t_1 \in T_1$, $\psi_2(t_2) = 1, \forall t_2 \in T_2$ and $\psi_{1,2}(\boldsymbol{t}) = \ln\left(\Pr_{\boldsymbol{t}'\sim D''}(\boldsymbol{t}' = \boldsymbol{t})\right)$. Since Distribution D'' samples from the product distribution D with probability $1 - \alpha$, we have that for each $\boldsymbol{t} \in T$, $\Pr_{\boldsymbol{t}'\sim D''}(\boldsymbol{t}' = \boldsymbol{t}) \ge (1 - \alpha)p^2$. This is true because with probability $1 - \alpha$ we sample from the product distribution D, and with probability at least p^2 we sample the type $\boldsymbol{t}$, therefore the MRF we just described has finite inverse temperature and $\Delta = \beta(\boldsymbol{t}) = \max_{\boldsymbol{t}\in\text{Supp}(D')} |\log\left(\Pr_{\boldsymbol{t}'\sim D'}[\boldsymbol{t}' = \boldsymbol{t}]\right)| \le |\log\left((1 - \alpha)p^2\right)|$. In the case where the distribution is over two items, $\beta(\boldsymbol{t}) = \max_{\boldsymbol{t}\in T} |\psi_{1,2}(\boldsymbol{t})|$. Moreover we can easily verify that the MRF we just described has the same joint distribution as D''. Therefore D'' can be modeled as a MRF with finite inverse temperature.

We now prove that $\text{Rev}(D'') \ge \alpha'\text{Rev}(D')$. This is true because we can simply use the optimal mechanism that induces $\text{Rev}(D')$ on D''. Since D'' take a sample from D' with probability α', we are guaranteed that this mechanism has revenue at least $\alpha'\text{Rev}(D')$ on D''.

The fact that D'' has the same marginal distributions as D' follows from the sampling procedure. □

We also need the following important result from [30].

Lemma 21 (Theorem A from [30])**.** *For any positive number N, there exists a two item correlated distribution D, such that for an additive buyer whose type is sampled from D, $\frac{\textsc{Rev}(D)}{\max\{\text{B}\textsc{Rev}(D), \text{S}\textsc{Rev}(D)\}} > N$.*

Equipped with Lemma 20 and 21, we are ready to prove Theorem 7.

Proof of Theorem 7: Let D' be the distribution over two items that is guaranteed to exist by Lemma 21. Since D' is a two dimensional distribution, its Dobrushin coefficient is at most 1.

Apply Lemma 20 to D' with parameter $\alpha' = \alpha$ to create another distribution D which has the same marginals as D' but with a Dobrushin coefficient at most α. Moreover, D can be expressed as a MRF with finite inverse temperature. Clearly, $\textsc{Rev}(D) \geq \alpha \cdot \textsc{Rev}(D')$, as one can simply achieve the RHS under distribution D using the optimal mechanism designed for D'. Also, $\text{S}\textsc{Rev}(D') = \text{S}\textsc{Rev}(D)$ as the two distributions have the same marginals. Finally, $\text{B}\textsc{Rev}(D') \leq 2\text{S}\textsc{Rev}(D')$. Suppose b is the optimal price for the bundle, then we can set the two items separately each at price $b/2$. Clearly, whenever the bundle is sold, at least one item is sold. To conclude, $\frac{\textsc{Rev}(D)}{\max\{\text{B}\textsc{Rev}(D), \text{S}\textsc{Rev}(D)\}} \geq \frac{\textsc{Rev}(D)}{2\text{S}\textsc{Rev}(D)} \geq \frac{\alpha \cdot \textsc{Rev}(D')}{2\text{S}\textsc{Rev}(D')} > \frac{\alpha}{2} \cdot N$.

□

7.2 Lower Bound for the Copies Setting

In this section, we show that if the analysis uses the optimal revenue in the COPIES setting as part of the benchmark for the optimal revenue in the original setting (as in our analysis), the exponential dependence on the maximum weighted degree Δ in the approximation ratio is unavoidable. Note that we also showed that the approximation ratio must have polynomial dependence on Δ no matter what approach is used (Theorem 3).

Theorem 8. *For any value of $n \in \mathbb{N}$ and $\beta \in \mathbb{R}_+$ there exists a type distribution D over $n+1$ items, such that D can be represented by a MRF with only pairwise potentials and maximum weighted degree $\Delta \leq \beta \cdot n$. Moreover, for an additive or unit-demand buyer, the expected optimal revenue in the COPIES settings w.r.t. D can be arbitrarily close to $\frac{1}{2}\exp(2\beta n)$, while* $\max\{\text{B}\textsc{Rev}, \text{S}\textsc{Rev}\} < 2$.

Proof of Theorem 8: We construct the MRF in the following way. The first item has support $T_1 = \{2^0, 2^1, 2^2, \ldots, 2^{k^n-1}\}$, where $k \in \mathbb{N}$ is going to be defined later. Let $\varepsilon_1, \ldots, \varepsilon_k$ be some tiny non-negative values, and the support of the other items' distributions is $R = \{\varepsilon_i\}_{i \in [k]}$. We consider the following node potential for the first item:

$$\psi_1(2^i) = \begin{cases} \ln(\frac{1}{2^{i+1}}) & \text{if } 0 \leq i \leq k^n - 2 \\ \ln(\frac{1}{2^i}) & \text{if } i = k^n - 1 \end{cases}$$

The node potentials for the other items is: $\psi_i(a) = \ln\left(\frac{1}{\exp(\beta)+(k-1)\exp(-\beta)}\right)$ for all $i \in [2, n+1]$ and $a \in R$.

Note that $|T_1| = k^n$ and $|R^n| = k^n$, therefore for each $t_1 \in T_1$, we can map it to a unique $t_{-1} \in R^n$. Formally, we consider a bijective function $c : T_1 \to R^n$.

We define pair-wise potentials between the first item and the j-th item:

$$\psi_{1,j}(2^i, \varepsilon_\ell) = \begin{cases} \beta & \text{if } \varepsilon_\ell = c(2^i)_j \\ -\beta & \text{if } \varepsilon_\ell \neq c(2^i)_j \end{cases}$$

It is easy to verify that $\Delta \le \beta \cdot n$ for the constructed MRF.

Let Z be the normalizing constant so that the MRF with potentials $\{\psi_i\}_{i\in[n+1]}, \{\psi_{1,i}\}_{2\le i\le n+1}$ is a valid distribution. That is $Z = \sum_{t\in \textsc{Supp}(D)} \prod_{i\in[n+1]} \exp(\psi_i(t_i)) \prod_{2\le i\le n+1} \exp(\psi_{1,i}(t_1,t_i))$. For any $t_1 \in T_1$ we have that: $\Pr_{t'\sim D}\left[t'_1 = t_1 \wedge t'_{-1} = c(t_1)\right] = \frac{1}{Z}\exp(\psi_1(t_1))\frac{\exp(n\beta)}{(\exp(\beta)+(k-1)\exp(-\beta))^n}$.

$$\begin{aligned}
&\Pr_{t'\sim D}\left[t'_1 = t_1 \wedge t'_{-1} \neq c(t_1)\right] \\
=&\frac{1}{Z}\exp(\psi_1(t_1))\frac{1}{(\exp(\beta)+(k-1)\exp(-\beta))^n}\sum_{\substack{t_{-1}\in T_{-1}:\\ t_{-1}\neq c(t_1)}}\prod_{i\in[2,n+1]}\exp\left(\psi_{1,i}(t_1,\{t_{-1}\}_i)\right) \\
=&\frac{1}{Z}\exp(\psi_1(t_1))\frac{1}{(\exp(\beta)+(k-1)\exp(-\beta))^n}\sum_{i\in[1,n]}\binom{n}{i}(k-1)^i\,(\exp(-\beta))^i\,(\exp(\beta))^{n-i} \\
=&\frac{1}{Z}\exp(\psi_1(t_1))\frac{(\exp(\beta)+(k-1)\exp(-\beta))^n-\exp(n\beta)}{(\exp(\beta)+(k-1)\exp(-\beta))^n}
\end{aligned}$$

Thus for any $t_1 \in T_1$, the marginal probability: $f_1(t_1) = \frac{1}{Z}\exp(\psi_1(t_1))$. Note that $Z = \sum_{t_1\in T_1}\exp(\psi_1(t_1)) = \sum_{i=0}^{k^n-2}\frac{1}{2^{i+1}} + \frac{1}{2^{k^n-1}} = 1$ and $f_1(t_1) = \exp(\psi_1(t_1))$. Therefore the marginal distribution of the first item is an Equal Revenue Distribution, which means that the revenue of any posted price mechanism for the first item, cannot be more than 1. Moreover, if we choose $\varepsilon_1,\ldots,\varepsilon_k$ to be sufficiently small so that $\max_{x\in R} \le \frac{1}{2n}$, then any posted price mechanisms for the rest n items has revenue less or equal than $\frac{1}{2}$. Thus $\textsc{SRev} < 2$.

Now we consider the following Mechanism in the copies settings. We first collect the values for all buyers except the first one t_{-1}, then let the first buyer decide whether she wants to purchase the item at price $c^{-1}(t_{-1})$. This is essentially Ronen's lookahead auction [40]. A lower bound on the revenue of this mechanism in the COPIES settings is:

$$\begin{aligned}
\sum_{t_1\in T_1} t_1 \Pr_{t'\sim D}\left[c^{-1}(t'_{-1}) = t_1 \wedge t'_1 = t_1\right] &= \sum_{t_1\in T_1} t_1\exp(\psi_1(t_1))\frac{\exp(n\beta)}{(\exp(\beta)+(k-1)\exp(-\beta))^n} \\
&=\left(\frac{1}{1+(k-1)\exp(-2\beta)}\right)^n\sum_{t_1\in T_1} t_1\exp(\psi_1(t_1)) \\
&\ge\left(\frac{1}{1+(k-1)\exp(-2\beta)}\right)^n\sum_{t_1\in T_1}\frac{1}{2} \\
&=\frac{1}{2}\left(\frac{1}{1+(k-1)\exp(-2\beta)}\right)^n|T_1| \\
&=\frac{1}{2}\left(\frac{k}{1+(k-1)\exp(-2\beta)}\right)^n
\end{aligned}$$

Where the last inequality follows from the definition of $\psi_1(t_1)$.

Note that if we fix β and n, and let $k\to\infty$, then $\lim_{k\to\infty}\left(\frac{k}{1+(k-1)\exp(-2\beta)}\right)^n = \exp(2\beta n)$.

Therefore as $k\to\infty$, the lower bound of the revenue of the proposed mechanisms becomes $\frac{\exp(2\beta n)}{2}$. Since we assumed that the value of the agent for each item except the first is less or equal than $\frac{1}{2n}$, then the value of the agent for all but the first item is less or equal than $\frac{1}{2}$. This implies that if the agent buys the whole bundle at price p, then she also buys the first item at price $p-\frac{1}{2}$. Let $\textsc{Rev}_1$ be the revenue of the posted price mechanism on the first item. Since the marginal of the fist item is the Equal Revenue Distribution, then $\textsc{Rev}_1 \le 1$. Moreover by the argument described above, we have that $\textsc{BRev} \le \textsc{Rev}_1 + \frac{1}{2} < 2$. Thus $\max\{\textsc{SRev},\textsc{BRev}\} < 2$. □

Acknowledgement. The authors would like to thank Jessie Huang for creating Figure 1 and to thank Costis Daskalakis for helpful discussions in the early stages of the paper.

References

[1] 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.

[2] 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.

[3] Saeed Alaei, Hu Fu, Nima Haghpanah, and Jason D. Hartline. The simple economics of approximately optimal auctions. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 628–637. IEEE Computer Society, 2013.

[4] Moshe Babaioff, Nicole Immorlica, Brendan Lucier, and S. Matthew Weinberg. A simple and approximately optimal mechanism for an additive buyer. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 21–30, 2014.

[5] MohammadHossein Bateni, Sina Dehghani, MohammadTaghi Hajiaghayi, and Saeed Seddighin. Revenue maximization for selling multiple correlated items. In *Algorithms-ESA 2015*, pages 95–105. Springer, 2015.

[6] Anand Bhalgat, Sreenivas Gollapudi, and Kamesh Munagala. Optimal auctions via the multiplicative weight method. In Michael J. Kearns, R. Preston McAfee, and Éva Tardos, editors, *ACM Conference on Electronic Commerce, EC '13, Philadelphia, PA, USA, June 16-20, 2013*, pages 73–90. ACM, 2013.

[7] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. A sharp concentration inequality with applications. *Random Struct. Algorithms*, 16(3):277–292, 2000.

[8] Patrick Briest, Shuchi Chawla, Robert Kleinberg, and S Matthew Weinberg. Pricing lotteries. *Journal of Economic Theory*, 156:144–174, 2015.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

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

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

[17] Yang Cai and Mingfei Zhao. Simple mechanisms for subadditive buyers via duality. In *the 49th Annual ACM Symposium on Theory of Computing (STOC)*, 2017.

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

[19] 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.

[20] Shuchi Chawla, David L. Malec, and Balasubramanian Sivan. The power of randomness in bayesian optimal mechanism design. In David C. Parkes, Chrysanthos Dellarocas, and Moshe Tennenholtz, editors, *Proceedings 11th ACM Conference on Electronic Commerce (EC-2010), Cambridge, Massachusetts, USA, June 7-11, 2010*, pages 149–158. ACM, 2010.

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

[22] Yuval Dagan, Constantinos Daskalakis, Nishanth Dikkala, and Siddhartha Jayanti. Learning from weakly dependent data under dobrushin's condition. In Alina Beygelzimer and Daniel Hsu, editors, *Conference on Learning Theory, COLT 2019, 25-28 June 2019, Phoenix, AZ, USA*, volume 99 of *Proceedings of Machine Learning Research*, pages 914–928. PMLR, 2019.

[23] Yuval Dagan, Constantinos Daskalakis, Nishanth Dikkala, and Siddhartha Jayanti. Learning from weakly dependent data under dobrushin's condition. *CoRR*, abs/1906.09247, 2019.

[24] Constantinos Daskalakis, Nikhil R. Devanur, and S. Matthew Weinberg. Revenue maximization and ex-post budget constraints. In Tim Roughgarden, Michal Feldman, and Michael Schwarz, editors, *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*, pages 433–447. ACM, 2015.

[25] 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.

[26] Constantinos Daskalakis, Nishanth Dikkala, and Gautam Kamath. Testing ising models. *IEEE Transactions on Information Theory*, 65(11):6829–6852, 2019.

[27] PL Dobruschin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory of Probability & Its Applications*, 13(2):197–224, 1968.

[28] RL Dobrushin and SB Shlosman. Completely analytical interactions: constructive description. *Journal of Statistical Physics*, 46(5-6):983–1014, 1987.

[29] Reza Gheissari, Eyal Lubetzky, Yuval Peres, et al. Concentration inequalities for polynomials of contracting ising models. *Electronic Communications in Probability*, 23, 2018.

[30] Sergiu Hart and Noam Nisan. Selling multiple correlated goods: Revenue maximization and menu-size complexity. *Journal of Economic Theory*, 183:991 – 1029, 2019.

[31] Nicole Immorlica, Sahil Singla, and Bo Waggoner. Prophet inequalities with linear correlations and augmentations. In Péter Biró, Jason Hartline, Michael Ostrovsky, and Ariel D. Procaccia, editors, *EC '20: The 21st ACM Conference on Economics and Computation, Virtual Event, Hungary, July 13-17, 2020*, pages 159–185. ACM, 2020.

[32] Ross Kindermann and Laurie Snell. *Markov random fields and their applications*, volume 1. American Mathematical Society, 1980.

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

[34] Ulrich Krengel and Louis Sucheston. On semiamarts, amarts, and processes with finite value. *Probability on Banach spaces*, 4:197–266, 1978.

[35] Steffen L. Lauritzen. *Graphical Models*. Oxford University Press, 1996.

[36] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. Markov chains and mixing times. *American Mathematical Soc., Providence*, 2009.

[37] Xinye Li and Andrew Chi-Chih Yao. On revenue maximization for selling multiple independently distributed items. *Proceedings of the National Academy of Sciences*, 110(28):11232–11237, 2013.

[38] Alexandros Psomas, Ariel Schvartzman, and S Matthew Weinberg. Smoothed analysis of multi-item auctions with correlated values. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 417–418, 2019.

[39] Dana Randall. Slow mixing of glauber dynamics via topological obstructions. In *Symposium on Discrete Algorithms: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, volume 22, pages 870–879, 2006.

[40] Amir Ronen. On approximating optimal auctions. In Michael P. Wellman and Yoav Shoham, editors, *Proceedings 3rd ACM Conference on Electronic Commerce (EC-2001), Tampa, Florida, USA, October 14-17, 2001*, pages 11–17. ACM, 2001.

[41] 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.

[42] Ester Samuel-Cahn et al. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *the Annals of Probability*, 12(4):1213–1216, 1984.

[43] Daniel W Stroock and Boguslaw Zegarlinski. The logarithmic sobolev inequality for discrete spin systems on a lattice. *Communications in Mathematical Physics*, 149(1):175–193, 1992.

[44] Liming Wu. Poincaré and transportation inequalities for gibbs measures under the dobrushin uniqueness condition. *Ann. Probab.*, 34(5):1960–1989, 09 2006.

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

A Missing Details of the Duality-base Benchmark

We provide the necessary information to derive the benchmark for XOS valuations in Appendix E. Deriving a benchmark for a constrained additive valuation is a simpler task. We summarize the benchmark for a constrained additive valuation in Definition 11.

Definition 11. *The duality framework provide the following bound:*

$$\begin{aligned}
\textsc{Rev}(M, v, D) &\leq \sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot (t_i \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] + \phi_i(t_i \mid t_{-i}) \cdot \mathbb{1}[\boldsymbol{t} \in R_i]) \\
&= \sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot \phi_i(t_i \mid t_{-i}) \cdot \mathbb{1}[\boldsymbol{t} \in R_i] \quad (\textsc{Single}) \\
&\quad + \sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot t_i \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] \, (\textsc{Non-Favorite})
\end{aligned}$$

Lemma 22. *We can bound* Non-Favorite *by* Core *and* Tail*. More specifically,*

$$\begin{aligned}
\textsc{Non-Favorite} &\leq \sum_{i \in [n]} \sum_{t_i > r} f_i(t_i) \cdot t_i \cdot \Pr_{\boldsymbol{t}' \sim D} \left[\boldsymbol{t}' \notin R_i \mid t'_i = t_i\right] \quad (\textsc{Tail}) \\
&\quad + \sum_{i \in [n]} \sum_{t_i \leq r} f_i(t_i) \cdot t_i \quad (\textsc{Core})
\end{aligned}$$

Proof.

$$\begin{aligned}
&\sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot t_i \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] \\
\leq &\sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f(\boldsymbol{t}) \cdot t_i \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] \\
= &\sum_{\boldsymbol{t} \in T} \sum_{i \in [n]} f_i(t_i) \cdot f_{-i}(t_{-i} \mid t_i) \cdot t_i \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] \\
= &\sum_{i \in [n]} \sum_{t_i \in T_i} f_i(t_i) \cdot t_i \sum_{t_{-i} \in T_{-i}} f_{-i}(t_{-i} \mid t_i) \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] \\
= &\sum_{i \in [n]} \sum_{t_i > r} f_i(t_i) \cdot t_i \sum_{t_{-i} \in T_{-i}} f_{-i}(t_{-i} \mid t_i) \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] \\
&\quad + \sum_{i \in [n]} \sum_{t_i \leq r} f_i(t_i) \cdot t_i \sum_{t_{-i} \in T_{-i}} f_{-i}(t_{-i} \mid t_i) \cdot \mathbb{1}[\boldsymbol{t} \notin R_i] \\
\leq &\sum_{i \in [n]} \sum_{t_i > r} f_i(t_i) \cdot t_i \cdot \Pr_{\boldsymbol{t}' \sim D} \left[\boldsymbol{t}' \notin R_i \mid t'_i = t_i\right] \quad (\textsc{Tail}) \\
&\quad + \sum_{i \in [n]} \sum_{t_i \leq r} f_i(t_i) \cdot t_i \quad (\textsc{Core})
\end{aligned}$$

□

B Missing Proofs from Section 4.2

Proof of Lemma 5:

For the term SINGLE we have that:

$$
\begin{aligned}
&\sum_{\boldsymbol{t}\in T}\sum_{i\in[n]} f(\boldsymbol{t})\cdot\pi_i(\boldsymbol{t})\cdot\phi_i(t_i \mid t_{-i})\cdot\mathbb{1}\left[\boldsymbol{t}\in R_i\right]\\
=&\sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i})\cdot\sum_{t_i\in T_i} f_i(t_i\mid t_{-i})\cdot\pi_i(\boldsymbol{t})\cdot\phi_i(t_i\mid t_{-i})\cdot\mathbb{1}\left[\boldsymbol{t}\in R_i\right]\\
\le&\sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i})\cdot\sum_{t_i\in T_i} f_i(t_i\mid t_{-i})\cdot\phi_i(t_i\mid t_{-i})^+\cdot\mathbb{1}\left[\boldsymbol{t}\in R_i\right]\\
=&\sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i})\cdot\sum_{t_i:\ (t_i,t_{-i})\in R_i} f_i(t_i\mid t_{-i})\cdot\phi_i(t_i\mid t_{-i})^+ \qquad (2)
\end{aligned}
$$

According to Definition 7, $\sum_{t_i:\ (t_i,t_{-i})\in R_i} f_i(t_i\mid t_{-i})\cdot\phi_i(t_i\mid t_{-i})^+ = \max_{p\ge\max_{j\neq i} t_j} p\cdot(1-F_i(p\mid t_{-i}))$, which is exactly the revenue of Ronen's auction when bidder i has the highest value and the other bidders have value t_{-i}. This completes the proof that SINGLE $\le$ RONEN$^{\text{COPIES}}$.

As the buyer has unit-demand valuation, the NON-FAVORITE term is at most the revenue of the second price auction in the COPIES settings. The second price auction can be viewed as a special case Ronen's mechanism, in which the price is always set to be the same as the second highest bid. Hence, NON-FAVORITE $\le$ RONEN$^{\text{COPIES}}$.

Recall that:

$$
\text{RONEN}^{\text{COPIES}} = \sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i})\cdot\max_{p\ge\max_{j\neq i} t_j} p\cdot(1-F_i(p\mid t_{-i}))
$$

Due to Lemma 2,

$$
\begin{aligned}
\text{RONEN}^{\text{COPIES}} &\le \sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i})\cdot\exp(4\Delta(\boldsymbol{t}))\max_{p\ge\max_{j\neq i} t_j} p\cdot(1-F_i(p))\\
&\le \sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i})\cdot\exp(4\Delta(\boldsymbol{t}))\sum_{t_i:\ (t_i,t_{-i})\in R_i} f_i(t_i)\cdot\phi_i(t_i)^+\\
&\le \exp(8\Delta(\boldsymbol{t}))\sum_{i\in[n]}\sum_{\boldsymbol{t}\in R_i} f(\boldsymbol{t})\cdot\phi_i(t_i)^+\\
&\le \exp(8\Delta(\boldsymbol{t}))\,\mathbb{E}_{\boldsymbol{t}}\left[\max_{i\in[n]}\phi_i(t_i)^+\right]
\end{aligned}
$$

□

Proof of Theorem 1: Let $\tau^* = \text{MEDIAN}_{\boldsymbol{t}}(\max_{i\in[n]}\phi_i(t_i)^+)$. We let price $p_i = \min\{p\in T_i : \phi_i(p)^+ \ge \tau^*\}$. We first provide a lower bound of the revenue of the posted-price mechanism under the prices $\{p_i\}_{i\in[n]}$.

For each item $i\in[n]$, let $\mathcal{E}_i$ denote the event $\{\boldsymbol{t}\in T : t_i\ge p_i\}$ and $\mathcal{E}_i'$ denote the event $\{\boldsymbol{t}\in T : t_j < p_j,\ \forall j\neq i\}$. Clearly, the buyer buys item i in event $\mathcal{E}_i\cap\mathcal{E}_i'$, so the revenue of the posted-price mechanism is at least

$$
\sum_{i\in[n]} p_i\cdot\Pr_{\boldsymbol{t}\sim D}\left[\mathcal{E}_i\cap\mathcal{E}_i'\right] \ge \exp(-4\Delta(\boldsymbol{t}))\cdot\sum_{i\in[n]} p_i\cdot\Pr_{\boldsymbol{t}\sim D}\left[\mathcal{E}_i\right]\Pr_{\boldsymbol{t}\sim D}\left[\mathcal{E}_i'\right].
$$

Note that $p_i \cdot \Pr_{\boldsymbol{t}\sim D}[\mathcal{E}_i] = \sum_{t_i \in T_i} f_i(t_i) \cdot \phi_i(t_i)^+ \cdot \mathbb{1}[\phi_i(t_i)^+ \geq \tau^*] = \tau^* \cdot \Pr_{t_i\sim D_i}[\phi_i(t_i)^+ \geq \tau^*] + \mathbb{E}_{t_i\sim D_i}[(\phi_i(t_i)^+ - \tau^*)^+]$ and $\Pr_{\boldsymbol{t}\sim D}[\mathcal{E}'_i] \geq 1/2$. Hence, the RHS of the inequality above is lower bounded by

$$\frac{\exp(-4\Delta(\boldsymbol{t}))}{2} \cdot \sum_{i\in[n]} \tau^* \cdot \Pr_{t_i\sim D_i}[\phi_i(t_i)^+ \geq \tau^*] + \mathbb{E}_{t_i\sim D_i}[(\phi_i(t_i)^+ - \tau^*)^+]$$
$$\geq \frac{\exp(-4\Delta(\boldsymbol{t}))}{2} \cdot \left(\frac{\tau^*}{2} + \sum_{i\in[n]} \mathbb{E}_{t_i\sim D_i}[(\phi_i(t_i)^+ - \tau^*)^+]\right)$$

The inequality is due to the union bound. By Lemma 3, the lower bound is at least $\frac{\exp(-4\Delta(\boldsymbol{t}))}{4} \cdot \mathbb{E}_{\boldsymbol{t}}\left[\max_{i\in[n]} \phi_i(t_i)^+\right]$. Combining this conclusion with Lemma 5, the revenue of the posted-price mechanism is at least $\frac{\textsc{Rev}(D)}{8\exp(12\Delta(\boldsymbol{t}))}$.

□

C Missing Proofs from Section 4.3

Proof of Lemma 6:

$$\begin{aligned}
\textsc{Single} &= \sum_{\boldsymbol{t}\in T}\sum_{i\in[n]} f(\boldsymbol{t}) \cdot \pi_i(\boldsymbol{t}) \cdot \phi_i(t_i \mid t_{-i}) \cdot \mathbb{1}[\boldsymbol{t} \in R_i] \\
&\leq \sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i}) \cdot \sum_{t_i:\ (t_i,t_{-i})\in R_i} f_i(t_i \mid t_{-i}) \cdot \phi_i(t_i \mid t_{-i})^+ \\
&= \sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i}) \cdot \max_{p\geq \max_{j\neq i} t_j} p \cdot (1 - F_i(p \mid t_{-i})) \\
&\leq \exp(4\Delta(\boldsymbol{t})) \cdot \sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i}) \cdot \max_{p\geq \max_{j\neq i} t_j} p \cdot (1 - F_i(p)) \\
&\leq \exp(4\Delta(\boldsymbol{t})) \cdot \sum_{i\in[n]}\sum_{t_{-i}\in T_{-i}} f_{-i}(t_{-i}) \cdot r_i \\
&\leq \exp(4\Delta(\boldsymbol{t})) \cdot r
\end{aligned}$$

The first equality is due to the definition of $\phi_i(t_i \mid t_{-i})$ (Definition 7). The second inequality follows from Lemma 2. The third and last inequalities follow from the definition of r_i and r.

Similarly, we can bound the term $\textsc{Tail} = \sum_{i\in[n]}\sum_{t_i>r} f_i(t_i) \cdot t_i \cdot \Pr_{\boldsymbol{t}'\sim D}[\boldsymbol{t}' \notin R_i \mid t'_i = t_i]$.

First, note that $\Pr_{\boldsymbol{t}'\sim D}[\boldsymbol{t}' \notin R_i \mid t'_i = t_i] \leq \Pr_{\boldsymbol{t}'\sim D}[\exists k \neq i : t'_k \geq t_i \mid t'_i = t_i]$. Therefore

$$\begin{aligned}
\textsc{Tail} \leq & \sum_{i\in[n]}\sum_{t_i>r} f_i(t_i)\cdot t_i \cdot \Pr_{\boldsymbol{t}'\sim D}\left[\exists k \neq i : t'_k \geq t_i \mid t'_i = t_i\right] \\
\leq & \exp(4\Delta(\boldsymbol{t}))\cdot \sum_{i\in[n]}\sum_{t_i>r} f_i(t_i)\cdot t_i \cdot \Pr_{t'_{-i}\sim D_{-i}}\left[\exists k \neq i : t'_k \geq t_i\right] \\
\leq & \exp(4\Delta(\boldsymbol{t}))\cdot \sum_{i\in[n]}\sum_{t_i>r} f_i(t_i)\cdot t_i \cdot \left(\sum_{k\neq i}\Pr_{t'_k\sim D_k}\left[t'_k \geq t_i\right]\right) \\
\leq & \exp(4\Delta(\boldsymbol{t}))\cdot \sum_{i\in[n]}\sum_{t_i>r} f_i(t_i)\cdot \sum_{k\neq i} r_k \\
\leq & \exp(4\Delta(\boldsymbol{t}))\cdot \sum_{i\in[n]} r\cdot \sum_{t_i>r} f_i(t_i) \\
\leq & \exp(4\Delta(\boldsymbol{t}))\cdot \sum_{i\in[n]} r_i \\
= & \exp(4\Delta(\boldsymbol{t}))\cdot r
\end{aligned}$$

The second inequality is due to Lemma 2. The third inequality follows from the union bound. The fourth and sixth inequalities hold because $r_k \geq t_i \cdot \Pr_{t'_k\sim D_k}[t'_k \geq t_i]$ and $r_i \geq r\cdot(1-F_i(r))$.

□

Proof of Lemma 7: We have that:

$$\sum_{\substack{c_i\leq r\\ c_j\leq r}} c_i c_j \Pr_{\boldsymbol{t}\sim D}[t_i = c_i \wedge t_j = c_j] \leq \exp(4\Delta(t))\cdot \sum_{\substack{c_i\leq r\\ c_j\leq r}} c_i f_i(c_i)\cdot c_j f_j(c_j) = \exp(4\Delta(t))\,\mathbb{E}[C_i]\,\mathbb{E}[C_j]$$

The inequality follows from Lemma 2. Therefore, $\textsc{Cov}[C_i, C_j] \leq (\exp(4\Delta(t)) - 1)\,\mathbb{E}[C_i]\,\mathbb{E}[C_j]$. Note that $\textsc{Var}[C] = \sum_{i\in[n]} \textsc{Var}[C_i^2] + \sum_{i\neq j}\textsc{Cov}(C_i, C_j) \leq \sum_{i\in[n]} \mathbb{E}[C_i^2] + \sum_{i\neq j}\textsc{Cov}(C_i, C_j)$. Using Lemma 9 from [14], we can bound $\sum_{i\in[n]} E[C_i^2]$ by $2r^2$. Hence,

$$\begin{aligned}
\textsc{Var}[C] \leq & 2r^2 + (\exp(4\Delta(\boldsymbol{t})) - 1)\sum_{i\neq j}\mathbb{E}[C_i]\,\mathbb{E}[C_j] \\
\leq & 2r^2 + (\exp(4\Delta(\boldsymbol{t})) - 1)\left(\sum_{i\in[n]}\mathbb{E}[C_i]\right)^2 \\
= & 2r^2 + (\exp(4\Delta(\boldsymbol{t})) - 1)\,\mathbb{E}[C]^2
\end{aligned}$$

□

Proof of Theorem 2: First we present the Paley-Zygmund inequality. For a non-negative random variable X, Paley-Zygmund inequality implies that for $\theta \in [0,1]$ we have that:

$$\Pr\left[X > \theta\,\mathbb{E}[X]\right] \geq (1-\theta)^2 \frac{1}{1 + \frac{\textsc{Var}[X]}{\mathbb{E}[X]^2}}$$

By the Paley-Zygmund inequality and Lemma 7, we derive the following inequality:

$$\Pr\left[C \geq \frac{\mathbb{E}[C]}{2}\right] \geq \frac{1}{4} \cdot \frac{1}{1 + \textsc{Var}[C]/\,\mathbb{E}[C]^2} \geq \frac{1}{4\left(\exp(4\Delta(\boldsymbol{t})) + 2r^2/\,\mathbb{E}[C]^2\right)}. \tag{3}$$

If $\mathbb{E}[C] \leq \sqrt{2}r$, then according to Lemma 6,

$$\textsc{Rev}(D) \leq \textsc{Core} + \textsc{Tail} + \textsc{Single} \leq \left(2\exp(4\Delta(\boldsymbol{t})) + \sqrt{2}\right) \cdot \textsc{SRev}.$$

Otherwise, Equation (3) implies that

$$\Pr_{\boldsymbol{t}\sim D}\left[\sum_{i\in[n]} t_i \geq \frac{\mathbb{E}[C]}{2}\right] \geq \Pr\left[C \geq \frac{\mathbb{E}[C]}{2}\right] \geq \frac{1}{4\left(\exp(4\Delta(\boldsymbol{t})) + 1\right)}.$$

Therefore, if we sell the grand bundle at price $\frac{\mathbb{E}[C]}{2} = \frac{\textsc{Core}}{2}$, it will be sold with probability at least $\frac{1}{4(\exp(4\Delta(\boldsymbol{t}))+1)}$. Thus $8\left(\exp(4\Delta(t)) + 1\right) \cdot \textsc{BRev} \geq \textsc{Core}$.

Combining everything, we have $\left(2\exp(4\Delta(\boldsymbol{t})) + \sqrt{2}\right) \cdot \textsc{SRev} + 8\left(\exp(4\Delta(t)) + 1\right) \cdot \textsc{BRev} \geq \textsc{Rev}(D)$. $\square$

D Missing Proofs from Section 5

Proof of Lemma 9:

We use r to denote SRev. We remind the readers that we use r to denote SRev, which is the revenue of the optimal posted price auction, in which we only allow the buyer to purchase at most one item.

We note that Single term in the XOS case, is the same as the the Single term in the Unit-Demand case in Section 4.2, if we consider that the buyer has valuation $V_i(\boldsymbol{t})$ for the i-th item. In section 4.2, using Lemma 5 we proved that $\textsc{Single} \leq \textsc{Ronen}^{\textsc{Copies}}$. Therefore it is enough to prove that there exists a posted price mechanism that allows the buyer to only pick her favorite item such that its revenue is at least $\frac{\textsc{Ronen}^{\textsc{Copies}}}{4\exp(12\Delta(\boldsymbol{t}))}$. A corollary of Theorem 1 is that there exists a posted price Mechanism M_p such that $\textsc{Ronen}^{\textsc{Copies}} \leq 4\exp(12\Delta(\boldsymbol{t})\textsc{Rev}(M_p)$, which concludes the proof for the term Single.

Next, we consider the term Tail. We remind the readers that $2r$ is a cutoff we use to separate the (Core) and the (Tail) term. The reason we chose this specific value is that we can bound the sum of the marginal probability that any item has value greater or equal than $2r$.

Lemma 23. *We have that:*

$$\sum_{i\in[n]} \Pr_{t_i\sim D_i}[V_i(t_i) \geq 2r] \leq \exp(4\Delta(\boldsymbol{t}))$$

Proof. We can lower bound $\Pr_{\boldsymbol{t}\sim D}[\exists i : V_i(t_i) \geq 2r]$ as the sum of the following disjoint events:

$$\begin{aligned}
&\Pr_{\boldsymbol{t}\sim D}[\exists i : V_i(t_i) \geq 2r] \\
\geq &\sum_{i\in[n]} \Pr_{\boldsymbol{t}\sim D}[V_i(t_i) \geq 2r \wedge \max_{j\neq i}\{V_j(t_j)\} < 2r]
\end{aligned}$$

Using Lemma 2 with sets $\mathcal{E} = \{t_i \in T_i : V_i(t_i) \geq 2r\}$ and $\mathcal{E}' = \{t_{-i} \in T_{-i} : \max_{j \neq i}\{V_j(t_j)\} < 2r\}$, we have that:

$$\begin{aligned}
&\sum_{i\in[n]} \Pr_{\boldsymbol{t}\sim D}[V_i(t_i) \geq 2r \wedge \max_{j\neq i}\{V_j(t_j)\} < 2r] \\
&\geq \sum_{i\in[n]} \Pr_{t_i\sim D_i}[V_i(t_i) \geq 2r] \exp(-4\Delta(\boldsymbol{t})) \Pr_{t\sim D}[\max_{j\neq i}\{V_j(t_j)\} < 2r] \\
&\geq \exp(-4\Delta(\boldsymbol{t})) \Pr_{t\sim D}[\max_{j}\{V_j(t_j)\} < 2r] \sum_{i\in[n]} \Pr_{t_i\sim D_i}[V_i(t_i) \geq 2r]
\end{aligned}$$

Note that $\Pr_{\boldsymbol{t}\sim D}[\exists i : V_i(t_i) \geq 2r] \leq \frac{1}{2}$. This is true because if we set the price of every item at $2r$, then if any item is bought with probability greater than $\frac{1}{2}$, we have revenue greater than r, a contradiction. Moreover $\Pr_{\boldsymbol{t}\sim D}[\max_j\{V_j(t_j)\} < 2r] = 1 - \Pr_{\boldsymbol{t}\sim D}[\exists i : V_i(t_i) \geq 2r] \geq \frac{1}{2}$. By these observations, we can conclude that:

$$\begin{aligned}
\frac{1}{2} &\geq \Pr_{\boldsymbol{t}\sim D}[\exists i : V_i(t_i) \geq 2r] \\
&\geq \exp(-4\Delta(\boldsymbol{t})) \Pr_{t\sim D}[\max_{j}\{V_j(t_j)\} < 2r] \sum_{i\in[n]} \Pr_{t_i\sim D_i}[V_i(t_i) \geq 2r] \\
&\geq \exp(-4\Delta(\boldsymbol{t}))\frac{1}{2} \sum_{i\in[n]} \Pr_{t_i\sim D_i}[V_i(t_i) \geq 2r],
\end{aligned}$$

which implies that $\sum_{i\in[n]} \Pr_{t_i\sim D_i}[V_i(t_i) \geq 2r] \leq \exp(4\Delta(\boldsymbol{t}))$.

□

Now we are going to bound the term Tail.

For any fixed $t_i \in T_i$, using Lemma 2 on sets $\mathcal{E} = \{t_i\}$ and $\mathcal{E}' = \{\boldsymbol{t}_{-i} \in \boldsymbol{T}_{-i} : \exists j \neq i, V_j(t'_j) \geq V_i(t'_i)\}$ we have that:

$$\begin{aligned}
\textsc{Tail} &= \sum_{i\in[n]} \sum_{\substack{t_i\in T_i \\ V_i(t_i)\geq 2r}} f(t_i) \cdot V_i(t_i) \Pr_{\boldsymbol{t}'\sim D}\left[\boldsymbol{t}' \notin R_i \mid t'_i = t_i\right] \\
&\leq \sum_{i\in[n]} \sum_{\substack{t_i\in T_i \\ V_i(t_i)\geq 2r}} f(t_i) \cdot V_i(t_i) \Pr_{\boldsymbol{t}'\sim D}[\exists j \neq i : V_j(t'_j) \geq V_i(t_i) \mid t'_i = t_i] \\
&\leq \sum_{i\in[n]} \sum_{\substack{t_i\in T_i \\ V_i(t_i)\geq 2r}} f(t_i) \cdot V_i(t_i) \exp(4\Delta(\boldsymbol{t})) \Pr_{\boldsymbol{t}'\sim D}[\exists j \neq i : V_j(t'_j) \geq V_i(t_i)]
\end{aligned}$$

We consider the mechanism that posts price $V_i(t_i)$ at each item except item i, and allows the buyer to get her favorite item. The expected revenue of this mechanisms is exactly $V_i(t_i)\Pr[\exists j \neq i : V_j(t_j) \geq V_i(t_i)]$, which is at most r. This implies that:

$$\begin{aligned}
&\exp(4\Delta(\boldsymbol{t}))\sum_{i\in[n]}\sum_{\substack{t_i\in T_i:\\ V_i(t_i)\ge 2r}} f(t_i)V_i(t_i)\Pr_{\boldsymbol{t}_{-i}\sim D_{-i}}\left[\exists j\neq i: V_j(t_j)\ge V_i(t_i)\right]\\
\le&\exp(4\Delta(\boldsymbol{t}))\sum_{i\in[n]}\sum_{\substack{t_i\in T_i:\\ V_i(t_i)\ge 2r}} f_i(t_i)r\\
=&\exp(4\Delta(\boldsymbol{t}))r\sum_{i\in[n]}\Pr_{t_i\sim D_i}\left[V_i(t_i)\ge 2r\right]\\
\le&\exp(4\Delta(\boldsymbol{t}))\exp(4\Delta(\boldsymbol{t}))r\\
=&\exp(8\Delta(\boldsymbol{t}))\cdot r
\end{aligned}$$

Where the last inequality follows from Lemma 23.

□

Proof of Lemma 14: Define $g_i(\boldsymbol{t}_{-i}) = v(\boldsymbol{t}_{-i}, C(\boldsymbol{t}_{-i}))$, where $C(\boldsymbol{t}_{-i}) = \{j : V_j(\boldsymbol{t}_{-i}) < 2\text{SRev}\}$. When $i \notin C(\boldsymbol{t})$, $g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i}) = 0$. When $i \in C(\boldsymbol{t})$, $C(\boldsymbol{t}) = C(\boldsymbol{t}_{-i}) \cup \{i\}$. Additionally, $g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i}) \ge 0$ and $g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i}) \le V_i(\boldsymbol{t}) \le 2\text{SRev}$. Since $v(\cdot,\cdot)$ is a XOS function, there exists non-negative numbers $\{x_\ell\}_{\ell\in C(\boldsymbol{t})}$ such that $g(\boldsymbol{t}) = \sum_{\ell\in C(\boldsymbol{t})} x_\ell$ and $g_i(\boldsymbol{t}_{-i}) \ge \sum_{\ell\in C(\boldsymbol{t}_{-i})} x_\ell$. Therefore, $\sum_{i\in[n]}(g(\boldsymbol{t}) - g_i(\boldsymbol{t}_{-i})) \le \{x_\ell\}_{\ell\in C(\boldsymbol{t})} = g(\boldsymbol{t})$.

Combining Lemma 11, 12, and the fact that $g(\cdot)$ is 2SRev-self-bounding, we derive the stated upper bound of the variance of $g(\boldsymbol{t})$.

□

E Missing Details of the Revenue Benchmark for a XOS Buyer

Similar to [17], we are going to apply the duality framework on a relaxed version of the valuation function.

Definition 12 (Relaxed Valuation (Definition 5 from [16]))**.** *We define the relaxed subadditive valuation $v^r(\boldsymbol{t}, S)$ the following way:*

$$v^r(\boldsymbol{t}, S) = \begin{cases} v(\boldsymbol{t}, S\backslash\{i\}) + V_i(t_i) & \textit{if } \boldsymbol{t}\in R_i \textit{ and } i \in S \\ v(\boldsymbol{t}, S) & \textit{Otherwise} \end{cases}$$

The reason that we consider the relaxed valuation function is because that v^r is "additive" across the favorite item and the rest of the items, and this "additivity" plays a crucial role in obtaining an analyzable dual. Due to the non-monotonicity of the optimal revenue in multi-item auctions, it is not clear that the optimal revenue w.r.t. the relaxed valuation is higher than the original optimal revenue. The following Lemma shows that the optimal revenue under v^r is not too much smaller than the original optimal revenue, so it suffices to apply the Cai-Devanur-Weinberg duality [14, 17] on the relaxed valuation v^r.

Lemma 24 (Lemma 2 from [16])**.** *We define by $\sigma_S(\boldsymbol{t})$ the probability that the buyer with type $\boldsymbol{t}$ receives exactly the set S in Mechanism M. Then:*

$$\text{Rev}(M, v, D) \le 2\text{Rev}(v^r, D) + 2\sum_{\boldsymbol{t}\in\boldsymbol{T}}\sum_{S\subseteq 2^{[n]}} f(\boldsymbol{t})\sigma_S(\boldsymbol{t})\left(v^r(\boldsymbol{t}, S) - v(\boldsymbol{t}, S)\right),$$

where $\text{REV}(v^r, D)$ *is the optimal revenue under the relaxed valuation* v^r

Now we are going to show how to upper bound the $\text{REV}(v^r, D)$ with terms similar to the ones we studied in the cases where the valuation function was additive.

Lemma 25 (Theorem 1 and Lemma 33 from [16])**.** *For a Mechanism* $M = (\sigma_S, p)$ *and a flow* $\boldsymbol{\lambda} : T \times T \to \mathbb{R}$ *that satisfied the partial specification (See Figure 3 from [16]), we have that:*

$$\text{REV}(M, v^r, D) \leq \sum_{\boldsymbol{t} \in T} f(\boldsymbol{t}) \sum_{S \subseteq 2^{[n]}} \sigma_S(\boldsymbol{t}) \Phi^r(\boldsymbol{t}, S)$$

Where $\Phi^r(\cdot, \cdot) : T \times 2^{[n]} \to \mathbb{R}$ *is the virtual valuation function, defined as:*

$$\Phi^r(\boldsymbol{t}, S) = \begin{cases} v(\boldsymbol{t}, S \backslash \{i\}) + V_i(t_i) - \frac{1}{f(\boldsymbol{t})} \sum_{\boldsymbol{t}' \in \boldsymbol{T}} \lambda(\boldsymbol{t}', \boldsymbol{t}) \left(V_i(t'_i) - V_i(t_i)\right) & \textit{if } \boldsymbol{t} \in R_i \textit{ and } i \in S \\ v(\boldsymbol{t}, S) & \textit{Otherwise} \end{cases}$$

For $\boldsymbol{t} \in R_i$*, we set* $\Psi_i^r(\boldsymbol{t}) = V_i(t_i) - \frac{1}{f(\boldsymbol{t})} \sum_{\boldsymbol{t}' \in \boldsymbol{T}} \lambda(\boldsymbol{t}', \boldsymbol{t}) \left(V_i(t'_i) - V_i(t_i)\right)$*. So we have that:*

$$\Phi^r(\boldsymbol{t}, S) \leq \begin{cases} v_i(t_i, S \backslash \{i\}) + \Psi_i^r(\boldsymbol{t}) & \textit{if } \boldsymbol{t} \in R_i \textit{ and } i \in S \\ v(\boldsymbol{t}, S) & \textit{Otherwise} \end{cases}$$

The following lemma provides a way to set a flow that satisfies the partial specifications (See Figure 3 from [16]).

Lemma 26 (Adapted Claim 1 from [16])**.** *There exists a flow that satisfies the partial specifications (See Figure 3 by [16]) such that:*

$$\Psi_i^r(\boldsymbol{t}) \leq \phi_i(V_i(t_i) \mid \boldsymbol{t}_{-i})$$

Where by $\phi_i(V_i(\boldsymbol{t}) \mid \boldsymbol{t}_{-i})$ *we denote the ironed virtual value of* $V_i(t_i)$*, when* t_i *is sampled from* $D_{i|\boldsymbol{t}_{-i}}$*.*

Proof. First we are going to describe how to set a flow that satisfies the partial specification requirements such that for $\boldsymbol{t} \in R_i$ it holds that $\Psi_i^r(\boldsymbol{t}) \leq \phi_i^N(V_i(t_i) \mid \boldsymbol{t}_{-i})$, where by $\phi_i^N(V_i(t_i) \mid \boldsymbol{t}_{-i})$ we denote the non-ironed virtual value of $V_i(t_i)$, when t_i is sampled from $D_{i|\boldsymbol{t}_{-i}}$. Then the way to set a flow that satisfies the partial specifications such that for $\boldsymbol{t} \in R_i$ it holds that $\Psi_i^r(\boldsymbol{t}) \leq \phi_i(V_i(t_i) \mid \boldsymbol{t}_{-i})$ is similar to the ironing procedure of Section 4 by [14].

For any two types $\boldsymbol{t}, \boldsymbol{t}'$, $\lambda(\boldsymbol{t}', \boldsymbol{t}) > 0$ only if there exists $i \in [n]$ such that $\boldsymbol{t}, \boldsymbol{t}' \in R_i$, $\boldsymbol{t}_{-i} = \boldsymbol{t}_{-i}$ and $t_i = \text{argmax}\{\hat{t}_i \in T_i : V_i(t'_i) > V_i(\hat{t}_i)\}$. Let $\boldsymbol{t}' \in R_i$, and $V = \max\{V_i(t_i) : V_i(t'_i) > V_i(t_i)\}$, we define $D(\boldsymbol{t}') = \{\boldsymbol{t} \in T : V_i(t_i) = V \wedge t_{-i} = t'_{-i}\}$. Note that $\lambda(\boldsymbol{t}', \boldsymbol{t}) > 0$ only if $\boldsymbol{t} \in D(\boldsymbol{t}')$. For any $\boldsymbol{t}' \in R_i$ and $\boldsymbol{t} \in D(\boldsymbol{t}') \cap R_i$, we set $\lambda(\boldsymbol{t}', \boldsymbol{t})$ to be equal to $\frac{f(\boldsymbol{t})}{\Pr_{t'_i \sim D_i}[V_i(t'_i) = V_i(t_i) \wedge \boldsymbol{t}'_{-i} = \boldsymbol{t}_{-i}]}$ fraction of the total in flow at $\boldsymbol{t}'$. We note that for any type $\boldsymbol{t}' \in T$, the sum of fractions of flows that it pushes to other types is at most one:

$$\sum_{\boldsymbol{t} \in D(\boldsymbol{t}')} \frac{f(\boldsymbol{t})}{\Pr_{t'_i \sim D_i}[V_i(t'_i) = V_i(t_i) \wedge \boldsymbol{t}'_{-i} = \boldsymbol{t}_{-i}]} = 1$$

Therefore if $\sum_{\boldsymbol{t}\in T}\lambda(\boldsymbol{t}',\boldsymbol{t})<1$, we can dump any remaining flow in the sink. It is clear that this flow satisfies the partial specifications.

Moreover for any $\boldsymbol{t}\in R_i$, the total in flow of $\boldsymbol{t}$ is:

$$\sum_{\boldsymbol{t}'\in T}\lambda(\boldsymbol{t}',\boldsymbol{t})=\frac{f(\boldsymbol{t})}{\Pr_{t'\sim D}[V_i(t'_i)=V_i(t_i)\wedge \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]}\Pr_{t'\sim D}[V_i(t'_i)>V_i(t_i)\wedge \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]$$

Therefore we have that:

$$\begin{aligned}\Psi_i^r(\boldsymbol{t})=&V_i(\boldsymbol{t})-\frac{1}{f(\boldsymbol{t})}\frac{f(\boldsymbol{t})}{\Pr_{t'\sim D}[\{V_i(t'_i)=V_i(t_i)\wedge \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]}\Pr_{t'\sim D}[V_i(t'_i)>V_i(t_i)\wedge \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]\big(V_i(t'_i)-V_i(t_i)\big)\\ =&V_i(\boldsymbol{t})-\frac{\Pr_{t'\sim D}[V_i(t'_i)>V_i(t_i)\wedge \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]}{\Pr_{t'\sim D}[V_i(t'_i)=V_i(t_i)\wedge \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]}\big(V_i(t'_i)-V_i(t_i)\big)\\ =&V_i(\boldsymbol{t})-\frac{\Pr_{t'\sim D}[V_i(t'_i)>V_i(t_i)\mid \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]}{\Pr_{t'\sim D}[V_i(t'_i)=V_i(t_i)\mid \boldsymbol{t}'_{-i}=\boldsymbol{t}_{-i}]}\big(V_i(t'_i)-V_i(t_i)\big)\end{aligned}$$

Therefore, $\Psi_i^r(\boldsymbol{t})$ is equal to the non-ironed virtual valuation of $V_i(t_i)$, when t_i is sampled from $D_{i|t_{-i}}$.

□

Combining Lemma 24, Lemma 25 and Lemma 26 we get the following lemma.

Lemma 27 (Adapted version of Theorem 2 from [17])**.**

$$\begin{aligned}\textsc{Rev}(M,v,D)\leq&2\sum_{\boldsymbol{t}\in T}f(\boldsymbol{t})\sum_{i\in[n]}\pi_i(\boldsymbol{t})\phi(V_i(t_i)\mid \boldsymbol{t}_{-i})\mathbb{1}[t\in R_i]\quad(\textsc{Single})\\ &+4\sum_{\boldsymbol{t}\in T}f(\boldsymbol{t})\sum_{i\in[n]}v(\boldsymbol{t},[n]\backslash\{i\})\mathbb{1}[\boldsymbol{t}\in R_i](\textsc{Non-Favorite})\end{aligned}$$

We further decompose the (Non-Favorite) term the following way:

We note that in the case where the buyer is XOS, we chose $2\boldsymbol{SRev}$ as the value that separates the (Core) and the (Tail) term. We sum up in the following lemma.

Proof of Lemma 8:

$$
\begin{aligned}
(\textsc{Non-Favorite}) \leq & \sum_{\boldsymbol{t}\in T} f(\boldsymbol{t}) \sum_{i\in[n]} v(t,[n]/i)\mathbb{1}[t\in R_i] \\
\leq & \sum_{\boldsymbol{t}\in T} f(\boldsymbol{t}) \left(v(t,C(\boldsymbol{t})) + \sum_{i\in[n]} V_i(t_i)\mathbb{1}[V_i(t_i)\geq 2r \wedge t\notin R_i] \right) \\
\leq & \sum_{\boldsymbol{t}\in T} f(\boldsymbol{t})\cdot v(t,C(\boldsymbol{t})) \\
& + \sum_{i\in[n]} \sum_{\substack{t_i\in T_i \\ V_i(t_i)\geq 2r}} \sum_{\boldsymbol{t_{-i}}\in T_{-i}} f((t_i,\boldsymbol{t}_{-i}))\cdot V_i(t_i)\mathbb{1}[(t_i,\boldsymbol{t}_{-i})\notin R_i] \\
\leq & \sum_{\boldsymbol{t}\in T} f(\boldsymbol{t})\cdot v(t,C(\boldsymbol{t})) \\
& + \sum_{i\in[n]} \sum_{\substack{t_i\in T_i \\ V_i(t_i)\geq 2r}} \sum_{\boldsymbol{t_{-i}}\in T_{-i}} f(\boldsymbol{t_i})f(\boldsymbol{t}_{-i}\mid t_i)\cdot V_i(t_i)\mathbb{1}[(t_i,\boldsymbol{t}_{-i})\notin R_i] \\
\leq & \sum_{\boldsymbol{t}\in T} f(\boldsymbol{t})\cdot v(t,C(\boldsymbol{t})) \\
& + \sum_{i\in[n]} \sum_{\substack{t_i\in T_i \\ V_i(t_i)\geq 2r}} f(t_i)\cdot V_i(t_i) \sum_{\boldsymbol{t_{-i}}\in T_{-i}} f(\boldsymbol{t}_{-i}\mid t_i)\mathbb{1}[(t_i,\boldsymbol{t}_{-i})\notin R_i] \\
\leq & \sum_{\boldsymbol{t}\in T} f(\boldsymbol{t})\cdot v(t,C(\boldsymbol{t}))(\textsc{Core}) \\
& + \sum_{i\in[n]} \sum_{\substack{t_i\in T_i \\ V_i(t_i)\geq 2r}} f(t_i)\cdot V_i(t_i) \Pr_{\boldsymbol{t'}\sim D}\left[\boldsymbol{t'}\notin R_i \mid t'_i = t_i\right] (\textsc{Tail})
\end{aligned}
$$

The claim follows from the inequality above and Lemma 27. □

F Lower Bound: Polynomial Dependence on Δ

In this section, we prove that for sufficiently large values $m\in\mathbb{N}$, there exists an type distribution represented by a MRF with maximum weighted degree $O(m)$, such that the optimal revenue is at least $\Omega(m^{1/7})$ times the maximum revenue achieved by simple mechanisms.

To prove this statement, we first modify the construction of Hart and Nisan [30], where they prove the following Theorem. We present a high-level idea of the proofs in this section.

Lemma 28 (Theorem C from [30])**.** *There exists a two item correlated distribution D and a constant $c>0$, such that for any $m\in\mathbb{N}$, when a buyer with additive valuation is sampled from D,* $\frac{\textsc{Rev}(D)}{\textsc{BRev}(D)}\geq c\cdot m^{1/7}$.

Their construction, relies on the following lemma.

Lemma 29 (Proposition 7.5. from [30])**.** *Let $\{g_i\}_{0\leq i\leq m}\in[0,1]^n$ and $\{y_i\}_{0\leq i\leq m}\in\mathbb{R}^n_+$ be two sequences of $m+1$ points, such that $g_0=(0,\ldots,0)$. For $i\geq 1$ we define:*

$$gap_i := \min_{0\leq j<i}(g_i-g_j)\cdot y_i$$

For any $m \in \mathbb{N}$, there exists a sequence $\{g_i\}_{0\le i\le m}$ in $[0,1]^2$ such that $g_0 = (0,0)$ and for each $1 \le i \le m$, $||g_i||_2 \le 1$. Moreover, if we set $y_i = g_i$ for all $0 \le i \le m$, then $gap_i = \Omega\left(i^{-6/7}\right)$.

Their construction is developed inductively, by placing points on "shells" of fixed radius. More specifically, in the N-th shell, whose radius is $\frac{\sum_{i=1}^{N} i^{-3/2}}{\sum_{i=1}^{\infty} i^{-3/2}}$, they place $N^{3/4}$ points so that the angle between any pair of points in the same shell is $\Omega(N^{-3/4})$. They observed that for all points (except g_0), $||g_i||_2 = \Theta(1)$, since $\sum_{i=1}^{\infty} i^{-3/2} = \Theta(1)$ and $\min_{1\le i\le m} ||g_i||_2 = ||g_1||_2 = \Theta(1)$. In their proof, they only needed a lower bound for each gap_i, but we also need an upper bound. Lemma 30 provides us with that bound.

Lemma 30. *In the construction of the set of points $\{g_i\}_{0\le i\le m}$ in Proposition 7.5 in [30], if:*

- *we place the first point of each shell in the same line that passes through the origin $(0,0)$*
- *for the N-th shell, whose radius is $\frac{\sum_{i=1}^{N} i^{-3/2}}{\sum_{i=1}^{\infty} i^{-3/2}}$, and any point i in that shell (except the first point of that shell), there exists another point $j < i$ in that shell such that the angle between them is $\Theta(N^{-3/4})$,*

then if we consider $\{y_i\}_{0\le i\le m} = \{g_i\}_{0\le i\le m}$, for each point i that is in the N-th shell, we have that $gap_i = \Theta(N^{-3/2}) = \Theta(i^{-6/7})$.

Proof. First we note that there is no restriction that prevents us from placing the first point of each shell in the same line. This is because different shells, have different radius so there is no way two points coincide. It is also trivial to ensure that the for point i in the N-th shell, which is not the first point in the shell, there exists a point $j < i$ in the N-th shell such that the angle between them is $\Theta(N^{-3/4})$. We note that the only assumption about the set of points that was made in the proof of Proposition 7.5 in [30] (Lemma 29), was that for the N-th shell, the points are placed in a semicircle of the radius we described above and that the angle between any pair of points is $\Omega(N^{-3/4})$. Therefore the result of Lemma 29 applies to this point set too.

Now we are going to prove that for the first point i^* in the N-th shell, it holds that $gap_{i^*} = \Theta(N^{-3/2})$. Let j^* be the first point of the $N-1$ shell and i^* be the first point of the N-th shell, then $gap_{i^*} \le (g_{i^*} - g_{j^*}) \cdot g_{i^*} = (||g_{i^*}||_2 - ||g_{j^*}||_2)||g_{i^*}||_2 = \frac{N^{-3/2}}{\sum_{i=1}^{\infty} i^{-3/2}}||g_{i^*}||_2 = O(N^{-3/2})$, where the first equality holds because point i^* and point j^* lie in the same line that passes through the origin. Since the results of Lemma 29 holds here, we have that $gap_{i^*} = \Theta(N^{-3/2})$.

Next, we deal with the case where point i is not the first point of the N-th shell. In the proof of Proposition 7.5 in [30], they noted that for two points that have angle θ between them, it holds that $\cos(\theta) = 1 - \Omega(\theta^2)$. We note that for $\theta < \frac{\pi}{2}$, we can similarly prove that $\cos(\theta) = 1 - \Theta(\theta^2)$ using the Taylor expansion of $\cos(\theta)$. For any point j that is not the first point in the shell, there is another point $i < j$ such that the two point have an angle $\theta = \Theta(N^{-3/4})$. Since $||g_j||_2 = \Theta(1)$, we can conclude that $gap_i \le (g_i - g_j) \cdot g_i = ||g_i||_2^2 - ||g_j||_2||g_i||_2 \cos(\theta) = ||g_i||_2^2\Theta(N^{-3/2}) = \Theta(N^{-3/2})$. Finally, since the N-th shell contains $N^{3/4}$ points, if point i belongs to the N-th shell, then $i = \Theta(N^{7/4})$. Hence, $gap_i = \Theta(i^{-6/7})$. □

Lemma 31 (Modified version of Theorem C from [30])**.** *For any sufficiently large $m \in \mathbb{N}$, there exists a two item correlated distribution D with support $T = \textsc{Supp}(D)$ and an absolute constant $C > 2$ such that $\inf_{\boldsymbol{t}\in\textsc{Supp}(D)} f(\boldsymbol{t}) \ge C^{-m}$. Moreover, when a buyer with additive valuation is sampled from D, then there exists another absolute constant $c > 0$ such that $\frac{\textsc{Rev}(D)}{\textsc{BRev}(D)} \ge c \cdot m^{1/7}$.*

Proof. Given any sequences $\{g_i\}_{0\leq i\leq m}$ and $\{y_i\}_{0\leq m}$ and a target value ε, Proposition 7.1 in [30] constructs the following distribution D: (i) Construct a sequence of positive numbers $\{t_i\}_{1\leq i\leq m}$ that increases fast enough, so that (a) $\xi_i := ||x_i||_1$ is increasing, where $x_i := \frac{t_i y_i}{gap_i}$ and (b) $\frac{t_{i+1}}{t_i} \geq \frac{1}{\varepsilon}$; (ii) The buyer has type x_i with probability $\frac{\xi_1}{\xi_i} - \frac{\xi_1}{\xi_{i+1}}$. Proposition 7.1 shows that for any choice of $\{t_i\}_{1\leq i\leq m}$ that satisfies property (a) and (b) in step (1) of the construction, the corresponding distribution D has $\frac{\textsc{Rev}(D)}{\textsc{BRev}(D)} \geq (1-\varepsilon)\sum_{i=1}^m \frac{gap_i}{||y_i||_1}$.

If we choose $\{g_i\}_{1\leq i\leq m}$ and $y_i = g_i$ for all $i \in [m]$ as in Lemma 30 and $\{t_i\}_{1\leq i\leq m}$ that satisfies property (a) and (b), it is not hard to verify that $\sum_{i=1}^m \frac{gap_i}{||y_i||_1} = \Omega(m^{1/7})$. Hence, the gap between $\textsc{Rev}(D)$ and $\textsc{BRev}(D)$ is as stated in the claim. The problem with this construction is that we cannot lower bound the probability that the rarest type shows up, as ξ_i and ξ_{i+1} can be very close to each other. To fix this issue, we modify the construction by replacing property (a) with a strengthened property (a*) $\frac{\xi_i}{\xi_{i-1}} \in [2, C]$ for all $i > 1$, where C is an absolute constant that will be determined later.

We first argue that if (a*) is satisfied, then the rarest type shows with sufficiently large probability. More specifically, type x_i shows up with probability

$$\frac{\xi_1}{\xi_i} - \frac{\xi_1}{\xi_{i+1}} \geq \frac{\xi_1}{\xi_{i+1}} \geq C^{-i}.$$

Next, we argue that for the sequences $\{g_i\}_{1\leq i\leq m}$ and $\{y_i\}_{1\leq i\leq m}$ as described in Lemma 30, there exists a sequence $\{t_i\}_{1\leq i\leq m}$ that satisfies (a*) and (b). Note that

$$\frac{\xi_i}{\xi_{i-1}} = \frac{t_i}{t_{i-1}} \frac{||g_i||_1}{||g_{i-1}||_1} \frac{gap_{i-1}}{gap_i}$$

By the definition of $\{g_i\}_{1\leq i\leq m}$, each point g_i is placed in a shell of radius at least $\frac{1}{\sum_{i=1}^{\infty} i^{-3/2}}$, so $||g_i||_2 = \Theta(1)$ and $\frac{||g_i||_2}{||g_{i-1}||_2} = \Theta(1)$. Since $g_i \in [0,1]^2$, $||g_i||_1 = \Theta(||g_i||_2)$, which implies that $\frac{||g_i||_1}{||g_{i-1}||_1} = \Theta(1)$. According to Lemma 30, $gap_i = \Theta(N^{-3/2})$ if i belongs to the N-th shell, so $\frac{gap_{i-1}}{gap_i} = \Theta(1)$. Hence, there exists two positive absolute constants C_1 and C_2 such that $C_1 \cdot \frac{t_i}{t_{i-1}} \leq \frac{\xi_i}{\xi_{i-1}} \leq C_2 \cdot \frac{t_i}{t_{i-1}}$. For the rest of the proof, we take ε to be 1/2. If we choose $\{t_i\}_{1\leq i\leq m}$ such that $\frac{t_i}{t_{i-1}}$ to be $\max\{2/C_1, 2\}$ for all $i > 1$, $\frac{\xi_i}{\xi_{i-1}} \in [2, C]$ for some absolute constant C. As the construction above satisfies both property (a*) and (b), we have

$$\frac{\textsc{Rev}(D)}{\textsc{BRev}(D)} \geq \frac{1}{2}\sum_{i=1}^m \frac{gap_i}{||y_i||_1} = \Omega(m^{1/7})$$

for the induced distribution D.

□

Proof of Theorem 3: By Lemma 31, there exists a type distribution D and constants $c > 0, c' > 0$ such that $\inf_{\boldsymbol{t}\in \textsc{Supp}(D)} \Pr_{\boldsymbol{t'}\sim D}[\boldsymbol{t'} = \boldsymbol{t}] \geq c^m$ and when a buyer has an additive valuation sampled from D, then $\frac{\textsc{Rev}(D)}{\textsc{BRev}(D)} \geq c' \cdot m^{1/7}$.

Using Lemma 20 on D with parameters $\alpha' = \frac{1}{2}$, we get a MRF D', such that its maximum weighted degree is bounded by $\Delta \leq \left|\log\left(\frac{c^{2m}}{2}\right)\right| = |2m \cdot \log(c) - \log(2)| = O(m)$ for sufficiently large $m > 0$. Moreover $\textsc{Rev}(D') \geq \frac{1}{2}\textsc{Rev}(D)$.

Since the marginal distributions of D and D' are the same, we have that $\text{SRev}(D') = \text{SRev}(D)$. We now prove that $\text{SRev}(D) \leq 2\text{BRev}(D)$. Let rev_1 be the optimal revenue when we only sell the first item, and rev_2 to be the optimal revenue when we only sell the second item. We can easily see that $\text{BRev} \geq \max(rev_1, rev_2)$. Since $\text{SRev} = rev_1 + rev_2$, we can conclude that $\text{SRev} \leq 2\text{BRev}$.

At this moment, we prove that $\text{BRev}(D) \leq 2\text{SRev}(D)$. Let p^* be the price induced by the optimal grand bundle mechanism. The revenue achieved by posting each item at price $p^*/2$ is a lower bound on $\text{SRev}(D)$. Moreover, if we sell each item at price $p^*/2$, then we are guaranteed to achieve at least half the revenue induced by $\text{BRev}(D)$ and our claim holds.

Thus $\frac{\text{Rev}(D')}{\max\{\text{BRev}(D'),\text{SRev}(D')\}} \geq \frac{1}{2}\frac{\text{Rev}(D')}{\text{SRev}(D')} \geq \frac{1}{4}\frac{\text{Rev}(D)}{\text{SRev}(D)} \geq \frac{1}{8}\frac{\text{Rev}(D)}{\text{BRev}(D)} \geq \frac{c \cdot m^{1/7}}{8}$, which concludes the proof. $\square$