Reverse Mechanism Design

Nima Haghpanah
Northwestern University
EECS Department
nima.haghpanah@gmail.com

Jason Hartline
Northwestern University
EECS Department
hartline@northwestern.edu

March 15, 2014

Abstract

Myerson's 1981 characterization of revenue-optimal auctions for single-dimensional agents follows from an amortized analysis of the incentives of a single agent. To optimize revenue in expectation, he maps values to *virtual values* which account for expected revenue gain but can be optimized pointwise. For single-dimensional agents the appropriate virtual values are unique and their closed form can be easily derived from *revenue equivalence*. A main challenge of generalizing the Myersonian approach to multi-dimensional agents is that the right amortization is not pinned down by revenue equivalence.

For multi-dimensional agents, the optimal mechanism may be very complex. Complex mechanisms are impractical and rarely employed. We give a framework for *reverse mechanism design*. Instead of solving for the optimal mechanism in general, we assume a (natural) specific form of the mechanism. As an example of the framework, for agents with unit-demand preferences, we restrict attention to mechanisms that sell each agent her favorite item or nothing. From this restricted form, we will derive multi-dimensional virtual values. These virtual values prove this form of mechanism is optimal for a large class of item-symmetric distributions over types. A corollary of this result is a sufficient condition for optimality of bundle pricing for general distributions. As another example of our framework, for bidders with additive preferences, we derive conditions for the optimality of posting a single price for the grand bundle.

1 Introduction

Optimal mechanisms for agents with multi-dimensional preferences are generally complex. This complexity makes them challenging to solve for and impractical to run. In a typical mechanism design approach, a model is posited and then the optimal mechanism is designed for the model. Successful mechanism design gives mechanisms that one could at least imagine running. By this measure, multi-dimensional mechanism design has had only limited success. In this paper we take the opposite approach, which we term *reverse mechanism design*. We start by imagining a restriction on mechanisms that would make them simple and reasonable to run, then we solve for sufficient conditions for a restricted mechanism to be optimal (among all mechanisms). Our approach is successful if the conditions under which restricted mechanisms are optimal are broad and representative of relevant settings.

This paper has two main contributions. The first is in codifying the *method of virtual values* from single-dimensional auction theory and extending it to agents with multi-dimensional preferences. The second is in applying this method to two paradigmatic classes of multi-dimensional preferences. The first class is unit-demand preferences (e.g., a homebuyer who wishes to buy at most one house); for this class we give sufficient conditions under which the posting a uniform price for all items is optimal. This result generalizes one of Alaei et al. (2013) for a consumer with values uniform on interval $[0, 1]$, and contrasts with an example of Thanassoulis (2004) for a consumer with values uniform on interval $[5, 6]$ where uniform pricing is not optimal. The second class is additive preferences, for this class we give sufficient conditions under which the posting a price for the grand bundle is optimal. This result generalizes a recent result of Hart and Nisan (2013) and relates to work of Armstrong (1999). These results for single-agent pricing problems described above generize naturally to multi-agent auction problems.

Myerson's (1981) characterization of revenue optimal auctions for single-dimensional agents is the cornerstone of modern auction theory and mechanism design. This characterization is successful in describing simple and practical mechanisms in simple environments where the agents preferences are independent and identically distributed according to a well-behaved distribution. In this case, the optimal auction is reserve price based. Myerson's characterization is also successful in describing the complex optimal mechanism for agents with preferences that are non-identically distributed or distributed according to an ill-behaved distribution. However, due to this complexity, the resulting mechanism has limited relevance. The consequence of our work is similar in that we characterize simple optimal mechanisms for well-behaved preferences; but distinct in that does not have consequences for ill-behaved preferences.

Myerson's approach is based on mapping agent values to appropriately defined *virtual values* and then optimizing the *virtual surplus*, i.e., the sum of the virtual values of agents served. Importantly, this approach replaces the gobal objective optimizing revenue in expectation over profiles of agent types to the pointwise objective of optimizing virtual surplus on each profile of agent types. Furthermore, virtual surplus maximization leads to simple and practical optimal mechanism in many environments. The simplicity of analysis by virtual values and of mechanisms resulting from optimizing virtual values has lead to a rich theory single-dimensional auction theory. Our multi-dimensional virtual values similarly give a pointwise objective and their optimization results in simple optimal mechanisms.

In the remainder of this section we review virtual values in mechanism design for single-dimensional agents, describe the challenges in identifying virtual values for multi-dimensional preferences, and describe how restricting the form of the optimal mechanism make it possible to solve for

multi-dimensional virtual values. To keep the exposition simple, we focus on designing mechanisms for a single-agent.

Optimal Single-dimensional Mechanisms for Revenue. Revenue is a challenging objective for mechanism design because, due to the agent's incentives, it cannot be optimized pointwise. To illustrate this fact, consider a single agent wishing to buy a single item with value for the item drawn uniformly from the $[0, 1]$ interval. If we post a take-it-or-leave-it price of 0.3 then the agent will buy when his value is greater than 0.3 and pay 0.3; if we post a price of 0.5 the agent will buy when his value is greater than 0.5 and pay 0.5. If our agent has value 0.3 then the first mechanism has the best possible revenue and the second mechanism has an inferior revenue of zero. If our agent has value 0.5 then the second mechanism has the best possible revenue of 0.5 and the first mechanism has an inferior revenue of 0.3. While we cannot rank these mechanisms by revenue pointwise, i.e., for all values the agent might possess, in expectation over the distribution from which the agent's value is drawn, the mechanisms can be ranked and revenue tradeoffs across possible agent values can be optimized. With a uniform value on $[0, 1]$, the agent buys in the first mechanism at a price of 0.3 with probability 0.7 for a revenue of 0.21, and in the second mechanism at a price of 0.5 with probability 0.5 for a revenue of 0.25.

Myerson (1981) solved for revenue optimal auctions for agents with single-dimensional preferences. He gives a definition of virtual values, claims that optimization of virtual surplus gives the optimal auction, and proves this claim by utilizing two properties of the virtual values of his definition. The concept of virtual values as the quantity to be optimized can be abstrated by explicitly requiring these two priperties.

Definition 1. A *virtual value function* maps values pointwise to virtual values satisfying two properties:

VV1. The pointwise optimization of virtual values is *incentive compatible*, i.e., there exist payments for this allocation rule that induce the agent to truthfully report his value.

VV2. The virtual values are an *amortization* of the objective in that any incentive compatible mechanism, the expected sum of the virtual values of the agents served is equal to the expected revenue.

The definition of virtual value functions gives a roadmap to identifying the optimal mechanism: find a virtual value function (that satisfies the two conditions) and run the mechanism that maximizes virtual value pointwise. The idenfitication of a virtual value function reduces the problem of optimization of the expected revenue (a global quantity) to the optimization of virtual surplus (a pointwise quantity).

In our single-dimensional example of the uniform distribution above, the appropriate virtual value function is $2v - 1$. Optimizing virtual value pointwise means serving the agent if his virtual value is positive and not serving him if his value is negative. For this virtual value function, the agent will be served if his value is least 0.5. This virtual value function satisfies the incentive compatibility condition (VV1): serving the agent if her value is above 0.5 is incentive compatible and the appropriate payment is 0.5 if the agent is served and zero otherwise. This virtual value function satisfies the amortization condition (VV2): e.g., for posting any price p, the the expected

virtual value of the agent served and the expected revenue are equal. The former can be calculated by $\int_p^1 (2v-1)\,dv = p\,(1-p)$ and the latter was calculated as above as $p\,(1-p)$.[1]

Bayesian optimal mechanism design has emerged as a fundamental challenge in stochastic optimization. Techniques for Bayesian mechanism design extend the reach of algorithmic approaches beyond worst case analysis and provide a rich framework for algorithm design. Bayesian mechanism design is inherently a multistage optimization. In the first stage a mechanism is optimized where the input to the optimization is the distribution on agent preferences. In subsequent stages, the mechanism determined in the first stage is run on inputs (agent preferences) drawn from the distribution. Hartline (2012) surveys this connection. The virtual-value-based approach greatly simplifies the task of finding and understanding optimal mechanisms for single-dimensional agents; developing a theory for multi-dimensional virtual values is a key step in understanding the general applicability of this fundamental approach to stochastic optimization.

The Challenge of Multi-dimensional Preferences. As described above, if a virtual value function that satisfies properties (VV1) and (VV2) can be identified then the optimal mechanism design problem is solved. For agents with single-dimensional preferences a function that satisfies the amortization property (VV2) is unique and can be derived by a simple exercise. Omitting the details: uniqueness follows because there is only one path on the line between any value and the origin (this is essentially an integration by parts). It then remains to check the incentive compatibility property (VV1), i.e., that pointwise optimization of virtual value is incentive compatible, which via standard characterizations of incentive compatible mechanisms is a simple task as well. Multi-dimensional preferences are challenging because the amortization property (VV2) does not uniquely pin down the virtual value function. Omitting the details: non-uniqueness follows because from any point in a multi-dimensional space there are many paths between the point and the origin. This difficulty has prevented the design of mechanisms for multi-dimensional agents that follows the virtual-value-based approach.

Multi-dimensional Virtual Values. To resolve the non-uniqueness of functions that satisfy the amortized revenue property (VV2) we consider additional constraints that the optimality of a restricted mechanism would place on virtual values.

We walk through this approach for the example of a unit-demand agent and the restriction to mechanisms that post a uniform price for each item. On one hand, under uniform pricing, the agent will always choose to buy his favorite item, or none of all values are below the price. On the other hand, a mechanism that optimizes a multi-dimensional virtual value (for a unit-demand agent) would serve the agent the item he has the highest positive virtual value for, or none if all virtual values are negative. Synthesizing these constraints, for virtual surplus maximization to imply optimality of uniform pricing the virtual value funciton must satisfy the following additional conditions:

R1. The virtual value for the favorite item corresponds to the single-dimensional virtual value for the distribution of the value for the favorite item (i.e.., the distribution of the maximum value), i.e., it is a *single-dimensional projection.*[2]

[1] This amortization also holds for randomized mechanisms but we omit the calculation from this discussion.

[2] This point follows because the mechanism has effectively projected the agent's multi-dimensional preference onto a single dimension. In this single dimension the unique function that satisfies the amortization property (VV2) is the one given by single-dimensional virtual values.

R2. Unless all virtual values are negative, the virtual value for the agent's favorite item is at least all other items' virtual values, i.e., they give a *consistent order* of the favorite item.

Any virtual value function that satisfies the restricted conditions (R1) and (R2) can be easily seen to satisfy the incentive compatibility requirement (VV1) via the same argument as for the single-dimensional virtual value described above.[3] It remains to show that such a virtual value function exists that satisfies the amortized revenue condition (VV2). In fact, the distributions over values for which such a function satisfies the amortized revenue condition are exactly the ones where mechanisms of the restricted form are optimal.

Two-dimensional preferences are a special case where our restriction on virtual values pins down a two-dimensional virtual value function uniquely. Specifically, as the virtual value for the favorite item is fixed by the restriction, only the virtual value for the other item must be determined. Essentially, we are left with a single-dimensional problem and in a single dimension the function that satisfies the amortized revenue condition are unique. Our task is then to give sufficient conditions so that the unique function that the amortization condition (VV2) and the single-dimensional projection restriction (R1) also satisfies the ordering restriction (R2), i.e., that the virtual value for the non-favorite items is not more than the virtual value for the favorite item unless both are negative.

For higher-dimensional preferences virtual value functions for the non-favorite items are not pinned down. Instead of deriving a formula for them, we generalize the sufficient condition from the two-dimensional case and prove that there exists a virtual value function that satisfies the above restriction and the amortized revenue condition.

In this paper we use the restriction of selling the agent his favorite item as an example of the general framework of reverse mechanism design. The techniques we develop can be similarly applied to other restrictions, and these extensions are an important topic for future work.

Sufficient Conditions for Optimality of Restricted Mechanisms. The opening example of this introduction of selling one of many houses to a homebuyer with value distributed independently, identically, and uniformly from $[0, 1]$ satisfies the conditions for optimality of the restriction to selling him only his favorite item. More generally, our framework identifies sufficient conditions for the distribution over agent values. Importantly these conditions allow for positive correlation between the agent's value for distinct items, a.k.a, affiliation. This correlation is natural when values correlate, e.g., with initial wealth.

We defer precise statement of the sufficient condition and instead describe a large class of distributions that it contains. Imagine drawing an agent's values in the following manner: a value v for the agents favorite item is drawn from a fixed distribution, the item that the agent considers favorite i^* is drawn uniformly at random, the ratios between the favorite value and each non-favorite values are drawn, independently of v, from any correlated distribution. Without the independence assumption, this definition is completely general for item-symmetric distributions (though, it only requires symmetry across the favorite item's value and identity); moreover, our general condition relaxes the independence assumption to a log-supermodularity condition.

As a very simple example, for the two-dimensional case, consider the distribution where the agent's favorite value v is uniform from interval $[0, 1]$ and his value for the non-favorite item is θv

[3]This statement assumes that the single-dimensional virtual values corresponding to the distribution of the agent's maximum value are monotone non-decreasing. A more sophisticated construction described in Section 6.3 removes this assumption.

with θ drawn uniformly from $[0, 1]$. The two dimensional virtual values that satisfy the amortization constraint (VV2) and the single-dimensional projection restriction (R1) pin down the possible virtual values as $2v - 1$ for the favorite item and $\theta(2v - 1)$ for the non-favorite item. For $\theta \in [0, 1]$ as defined, these virtual values satisfy the ordering restriction (R2). Therefore, maximizing these virtual values gives an optimal mechanism.

The virtual values that our method gives for i.i.d. distributions do not generally satisfy the ordering restriction (e.g., when values are i.i.d. $U[5, 6]$ the optimal mechanism does not sell the agent his favorite item); however, we show that the ordering restriction is met for values drawn i.i.d. from any distribution that satisfies a geometric convexity property.

As another example of the application of our framework, we provide sufficient conditions for optimality of pricing the grand bundle for agents with additive preferences. Roughly speaking, for the case of two items, the condition states that the density function is a log-submodular function of the following two parameters: the value for the grand bundle, and the ratio between the value for the non-favorite and favorite items.

Related Work. The starting point of work in multi-dimensional optimal mechanism design the observation that an agent's utility must be a convex function of his private type (e.g., Rochet, 1985, cf. the envelope theorem). The second step is in writing revenue as the difference between the surplus of the mechanism and the agent's utility (e.g., Armstrong, 1996). The surplus is expressed can be expressed in terms of the gradient of the utility. The third step is in rewriting the objective in terms of either the utility (e.g., Wang and Tang, 2013; Hart and Nisan, 2013) or in terms of the gradient of the utility (e.g., Alaei et al., 2013; Daskalakis et al., 2013). This manipulation follows from an integration by parts. These works perform the integration by parts independently in each dimension. In our approach, in contrast, the integration by parts is performed in general and dependent on the form of the mechanism we wish to show is optimal.

Alaei et al. (2013) show that optimal multi-agent mechanism design can be reduced to optimal single-agent mechanism design by the construction of a single-dimensional virtual value (that satisfies similar properties to Definition 1) when the single-agent mechanism design problems satisfy a *revenue linearity* property. They prove that a unit-demand agent with values that are independently, identically, and uniformly distributed on the $[0, 1]$ interval is revenue linear. Our work extends the reach of this work as our multi-dimensional virtual value construction constitutes a proof of revenue linearity. Therefore, all of our optimal single-agent mechanisms automatically generalize to give optimal multi-agent mechanisms for the service constrained environments of Alaei et al. (2013).

Finally, there has been work looking at properties of single-agent problems that are sufficient for optimal mechanisms to have nice form. For instance Pavlov (2011) gives conditions the optimal mechanism for a multi-dimensional agent with additive utility to be deterministic; however, he does not describe optimal mechanisms in more detail. (Recall: an example from Thanassoulis (2004) shows that the optimal mechanism for an unit-demand agent with values for two items each drawn uniformly from $[5, 6]$ is randomized.) Hart and Nisan (2013) ask a very similar question to ours (among other questions) but for an agent with additive values for two items and the restriction that the optimal mechanism is to only set a price for the bundle. For this simple case, they derive conditions on the distribution of values that are sufficient for the restricted form to be optimal. Daskalakis et al. (2013) consider selling two items to agents with additive values for bundles of the item. Their work describes closed form solutions for some special cases of independent distributions.

Notably, their approach starts from the characterization of incentive compatibility by convexity of the agent's utility and the connection between incentive constraints and paths between types and the origin (Manelli and Vincent, 2006).

2 Preliminaries

2.1 The Setting

An agent is specified by a bounded set of possible types T normalized to be $T = [0,1]^m$, where each $\mathbf{t} = (t_1, \ldots, t_m) \in T$ is an m-dimensional vector of values for m items[4]. The type of the agent is drawn from a known distribution f. We do not require that the values for items be drawn independently. Agents are unit demand, and therefore allocations $\mathbf{x} \in [0,1]^m$ satisfy $\sum_i x_i \leq 1$. The utility of the agent with type $\mathbf{t}$ for allocation $\mathbf{x} \in [0,1]^m$ and payment $p \in \mathbb{R}$ is $\mathbf{t} \cdot \mathbf{x} - p$.

A single-agent mechanism is a pair of functions, the allocation function $\mathbf{x} : T \to [0,1]^m$ and the payment function $p : [0,1]^m \to \mathbb{R}$. A mechanism is *individually rational* if the utility of every type of the agent is at least zero,

$$\mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - p(\mathbf{t}) \geq 0, \qquad \forall \mathbf{t}.$$

A mechanism is *incentive compatible* if no type of the agent increases his utility by misreporting,

$$\mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - p(\mathbf{t}) \geq \mathbf{t} \cdot \mathbf{x}(\hat{\mathbf{t}}) - p(\hat{\mathbf{t}}), \qquad \forall \mathbf{t}, \hat{\mathbf{t}}.$$

2.2 Multivariable Calculus Notation

For a function $h : \mathbb{R}^k \to \mathbb{R}$, we use $\partial_j h : \mathbb{R}^k \to \mathbb{R}$ to denote the *partial derivative* of function h with respect to its j'th variable. The *gradient* of h is a vector field, denoted by $\nabla h : \mathbb{R}^k \to \mathbb{R}^k$, defined to be $\nabla h = (\partial_1 h, \ldots, \partial_k h)$. The *divergence* of a vector field $\boldsymbol{\alpha} : \mathbb{R}^k \to \mathbb{R}^k$ is denoted by $\nabla \cdot \boldsymbol{\alpha} : \mathbb{R}^k \to \mathbb{R}$ and is defined to be

$$\nabla \cdot \boldsymbol{\alpha} = \partial_1 \alpha_1 + \ldots + \partial_k \alpha_k.$$

We denote the integral of function $h : \mathbb{R}^k \to \mathbb{R}$ over a subset T of $\mathbb{R}^k$

$$\int_{\mathbf{t} \in T} h(\mathbf{t}) \, \mathrm{d}\mathbf{t}.$$

Let ∂T be the boundary of set T. We can now define the multi-variable *integration by parts* for functions $h : \mathbb{R}^k \to \mathbb{R}$ and $\boldsymbol{\alpha} : \mathbb{R}^k \to \mathbb{R}^k$ as follows

$$\int_{\mathbf{t} \in T} \nabla h(\mathbf{t}) \cdot \boldsymbol{\alpha}(\mathbf{t}) \, \mathrm{d}\mathbf{t} = \int_{\mathbf{t} \in \partial T} h(\mathbf{t})(\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\mathbf{t}) \, \mathrm{d}\mathbf{t} - \int_{\mathbf{t} \in T} h(\mathbf{t}) \nabla \cdot \boldsymbol{\alpha}(\mathbf{t}) \, \mathrm{d}\mathbf{t},$$

[4]Throughout the paper we maintain the convention of denoting a vector $\mathbf{v}$ by a bold symbol and each of its components v_i by a non-bold symbol

where $\boldsymbol{\eta}(\mathbf{t})$ is the outward-pointing unit normal vector of T at point $\mathbf{t}$. Setting h to be the constant function equal to 1 everywhere gives us the *divergence theorem*

$$\int_{\mathbf{t}\in T} \nabla \cdot \boldsymbol{\alpha}(\mathbf{t})\ \mathrm{d}\mathbf{t} = \int_{\mathbf{t}\in \partial T} (\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\mathbf{t})\ \mathrm{d}\mathbf{t}.$$

When $k = 1$, the integration by parts has the familiar form of

$$\int_{x=a}^{b} h'(x)g(x)\ \mathrm{d}x = h(x)g(x)\Big|_{x=a}^{b} - \int_{x=a}^{b} h(x)g'(x)\ \mathrm{d}x.$$

And the divergence theorem reduces to the *fundamental theorem of calculus*.

$$\int_{x=a}^{b} h'(x)\ \mathrm{d}x = h(b) - h(a).$$

2.3 Problem Formulation

A single agent mechanism $(\mathbf{x}, p)$ defines a utility function $u(\mathbf{t}) = \mathbf{t}\cdot\mathbf{x}(\mathbf{t}) - p(\mathbf{t})$. The following lemma connects the utility function of an IC mechanism with its allocation function.

Lemma 1 (Manelli and Vincent (2006))**.** *Function u is the utility function of an agent in an incentive compatible mechanism if and only if u is convex, and in that case, the agent's allocation is $\mathbf{x}(\mathbf{t}) = \nabla u(\mathbf{t})$.*

Notice that the payment function can be defined using the utility function and the allocation function as $p(\mathbf{t}) = \mathbf{t}\cdot\mathbf{x}(\mathbf{t}) - u(\mathbf{t})$. Using the above lemma, we can write payment to be $p(\mathbf{t}) = \mathbf{t}\cdot\nabla u(\mathbf{t}) - u(\mathbf{t})$. This allows us to write the following optimization problem as our revenue maximization problem, which is the starting point of the analysis of this paper.

$$\begin{aligned} \max_{\bar{\mathbf{x}},u} \quad & \int_{\mathbf{t}} [\mathbf{t}\cdot\nabla u(\mathbf{t}) - u(\mathbf{t})] f(\mathbf{t})\ \mathrm{d}\mathbf{t} \\ & \mathbf{x} = \nabla u; u \text{ is convex}, \\ & \forall \mathbf{t}, \bar{\mathbf{x}}(\mathbf{t}) \text{ is feasible allocation.} \end{aligned} \tag{1}$$

Notice that when $m = 1$, the above program is equivalent to the following familiar form of

$$\begin{aligned} \max_{\bar{\mathbf{x}}} \quad & \int_{v} [v\mathbf{x}(v) - \int_{z\le v} \mathbf{x}(z)dz] f(v)\ \mathrm{d}v \\ & \mathbf{x} \text{ is monotone non-decreasing}, \\ & \forall v, \bar{\mathbf{x}}(v) \le 1. \end{aligned}$$

3 Virtual Values via An Amortized Analysis

This section introduces our high level framework for proving optimality of auctions. As stated before, our analysis is based on Program 4 as formulation of the optimal auction design problem. Notice that the representation of the objective function includes both functions $u(\mathbf{t})$ and its gradient $\nabla u(\mathbf{t})$. It is more convenient to rewrite the objective in terms of only ∇u. The following lemma suggests a way to do so using integration by parts.

Lemma 2. *Let functions $u : \mathbb{R}^m \to \mathbb{R}$, $\boldsymbol{\alpha} : \mathbb{R}^m \to \mathbb{R}^m$, and $f : \mathbb{R}^m \to \mathbb{R}$ be such that for any $\boldsymbol{t} \in \mathbb{R}^m$, we have $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}) = -f(\boldsymbol{t})$, and that for all $\boldsymbol{t} \in \partial T$, $\boldsymbol{t} \neq 0$, we have $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) = 0$, where $\boldsymbol{\eta}(\boldsymbol{t})$ is the normal vector at $\boldsymbol{t}$. Then, if $u(0) = 0$,*

$$\int_{\boldsymbol{t} \in T} u(\boldsymbol{t}) f(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} = \int_{\boldsymbol{t} \in T} \nabla u(\boldsymbol{t}) \boldsymbol{\alpha}(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}.$$

Proof. Using integration by parts,

$$\int_{\mathbf{t} \in T} \nabla u(\mathbf{t}) \boldsymbol{\alpha}(\mathbf{t}) \, \mathrm{d}\mathbf{t} = \int_{\mathbf{t} \in \partial T} u(\mathbf{t}) (\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\mathbf{t}) \, \mathrm{d}\mathbf{t} + \int_{\mathbf{t} \in T} u(\mathbf{t}) f(\mathbf{t}) \, \mathrm{d}\mathbf{t},$$

where ∂T is the boundary of T. Notice that function inside the first integral on the right hand side is zero everywhere given the assumptions of the lemma, and therefore it integrates to zero. This gives us the desired result. □

Notice that any optimal mechanism gives zero utility to type $\mathbf{t} = 0$, that is $u(0) = 0$. Directly from the above analysis, and recalling that $\mathbf{x} = \nabla u$ for incentive compatible mechanisms, we conclude the following lemma.

Lemma 3. *Let $\boldsymbol{\alpha}$ be any function satisfying the conditions of Lemma 2, then the revenue of any incentive compatible singe-agent allocation $\boldsymbol{x}$ is*

$$\int_{\boldsymbol{t}} \boldsymbol{x}(\boldsymbol{t}) \cdot (\boldsymbol{t} f - \boldsymbol{\alpha}(\boldsymbol{t})) \, \mathrm{d}\boldsymbol{t}.$$

We call $\boldsymbol{\alpha}$ satisfying conditions of Lemma 2 an amortization of the objective (see Property (VV2) in Section 1).

For simplicity in the rest of the paper we assume that $\boldsymbol{\phi}(\mathbf{t}) = \mathbf{t} - \frac{\boldsymbol{\alpha}(\mathbf{t})}{f(\mathbf{t})}$, and rewrite revenue of a single agent mechanism as

$$\int_{\mathbf{t}} \mathbf{x}(\mathbf{t}) \cdot \boldsymbol{\phi}(\mathbf{t}) f(\mathbf{t}) \, \mathrm{d}\mathbf{t}.$$

We say that a mechanism $\mathbf{x}$ optimizes an objective $\boldsymbol{\phi}$ *ex-post*, if for any input $\mathbf{t}$ it selects an outcome that maximizes $\mathbf{x}(\mathbf{t}) \cdot \boldsymbol{\phi}(\mathbf{t})$. Our above analysis proves the following.

Theorem 4 (Meta Theorem)**.** *A mechanism $\boldsymbol{x}$ is optimal if there exists an* ironed virtual value function $\bar{\boldsymbol{\phi}}$ *and $\boldsymbol{\alpha}$ satisfying conditions of Lemma 2 such that*

1. *Any incentive compatible mechanism satisfies*

$$\int_{\boldsymbol{t}} \boldsymbol{x}(\boldsymbol{t}) \cdot (\boldsymbol{t}f - \boldsymbol{\alpha}(\boldsymbol{t})) \, \mathrm{d}\boldsymbol{t} \leq \int_{\boldsymbol{t}} \boldsymbol{x}(\boldsymbol{t}) \cdot \bar{\boldsymbol{\phi}}(\boldsymbol{t}) f(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}.$$

2. *Allocation $\boldsymbol{x}$ ex-post optimizes $\bar{\boldsymbol{\phi}}$.*

3. *The revenue of $\boldsymbol{x}$ is to its expected virtual value, that is,*

$$\int_{\boldsymbol{t}} \boldsymbol{x}(\boldsymbol{t}) \cdot (\boldsymbol{t}f - \boldsymbol{\alpha}(\boldsymbol{t})) \, \mathrm{d}\boldsymbol{t} = \int_{\boldsymbol{t}} \boldsymbol{x}(\boldsymbol{t}) \cdot \bar{\boldsymbol{\phi}}(\boldsymbol{t}) f(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}.$$

The property that ex-post optimization of $\boldsymbol{\alpha}$ is incentive compatible is equivalent to Property (VV1) in Section 1. We say that $\boldsymbol{\alpha}$ is (approximate) *certificate* for optimality of allocation rule $\mathbf{x}$, if $\boldsymbol{\alpha}$ and $\mathbf{x}$ satisfy the conditions of the above theorem.

Notice that if $\bar{\boldsymbol{\phi}} f = \mathbf{t}f - \boldsymbol{\alpha}$, the first and third conditions of Theorem 4 are automatically satisfied, which leads us to the following corollary of Theorem 4. We sometimes prove the optimality of auctions by applying the following theorem first, and then prove stronger results by applying Theorem 4.

Theorem 5 (Unironed Meta Theorem)**.** *A mechanism $\boldsymbol{x}$ is optimal if there exists $\boldsymbol{\alpha}$ satisfying conditions of Lemma 2 such that $\boldsymbol{x}$ ex-post optimizes the* un-ironed virtual value function $\boldsymbol{\phi} := \boldsymbol{t} - \frac{\boldsymbol{\alpha}}{f}$.

4 Optimality of Item Pricing for Unit Demand Preferences

In this section we study the existence of proofs for uniform pricing systems for symmetric distributions in the unit demand setting. To simplify exposition we focus on the case of 2 items. The general case is only notationally more complicated and presented in Appendix A. We assume that $T = \{\mathbf{t} = (t_1, t_2) \in [0,1]^2 | t_1 \geq t_2\}$ as by symmetry the other half is similar.

We start by constructing a proof of optimality by applying Theorem 5. We will later improve our conditions by construing a proof using Theorem 4. Recall that we must design $\boldsymbol{\alpha}(\mathbf{t}) = (\mathbf{t} - \boldsymbol{\phi}(\mathbf{t}))f(\mathbf{t})$ so that $\nabla \cdot \boldsymbol{\alpha} = -f$, and the boundary conditions of Lemma 2 are satisfied. Rather than defining $\boldsymbol{\alpha}$ and $\bar{\boldsymbol{\phi}}$ directly, we find it more constructive to explain how it is derived to satisfy the desired properties. Our derivation assumes $\boldsymbol{\alpha}$ exists and uses its properties as a guideline to find the right form for it. In particular, we make an assumption that $\boldsymbol{\phi}_1$ is only a function of t_1, which is an important property when studying ex-post optimization of $\boldsymbol{\phi}$. This, together with the conditions of the divergence theorem, defines $\boldsymbol{\alpha}_1$. The property that $\nabla \cdot \boldsymbol{\alpha} = -f$ then implies a differential equation on $\boldsymbol{\alpha}_2$, which together with the boundary conditions of $\boldsymbol{\alpha}_2$ pins down the definition of $\boldsymbol{\alpha}_2$. We verify that the defined $\boldsymbol{\alpha}$ indeed satisfies the properties of Lemma 2, and then proceed to verify the structural properties that $\boldsymbol{\alpha}$ has to satisfy such that its ex-post optimization results in an incentive compatible allocation.

The first task is to define $\boldsymbol{\alpha}_1 = (t_1 - \boldsymbol{\phi}_1)f$. Since we require the expectation of $\mathbf{x} \cdot (\mathbf{t}f - \boldsymbol{\alpha})$ to be equal to the revenue of any mechanism $\mathbf{x}$, the equation must hold for any $\mathbf{x}_p$ that corresponds to posting a price p, for any p. Define $R(p)$ to be the revenue of posting a price p. Formally, $R(p) = p \int_{T(p)} f$, where $T(p) = \{\mathbf{t} \in T, t_1 \geq p\}$. The requirement that the expected virtual value of $\mathbf{x}_p$ is equal to $R(p)$ can be written as

$$\begin{aligned} R(p) &= \int_{\mathbf{t}\in T} (\mathbf{x}\cdot\boldsymbol{\phi} f)(\mathbf{t})\ \mathrm{d}\mathbf{t} \\ &= \int_{t_1\geq 0}\int_{t_2\leq t_1} \mathbf{x}(t_1,t_2)\cdot\boldsymbol{\phi}(t_1,t_2) f(t_1,t_2)\ \mathrm{d}t_2\ \mathrm{d}t_1 \\ &= \int_{t_1\geq p}\int_{t_2\leq t_1} \boldsymbol{\phi}_1(t_1,t_2) f(t_1,t_2)\ \mathrm{d}t_2\ \mathrm{d}t_1. \end{aligned}$$

Since we required that $\boldsymbol{\phi}_1(t_1,t_2) = \boldsymbol{\phi}_1(t_1)$, this becomes

$$\begin{aligned} R(p) &= \int_{t_1\geq p}\int_{t_2\leq t_1} \boldsymbol{\phi}_1(t_1) f(t_1,t_2)\ \mathrm{d}t_2\ \mathrm{d}t_1 \\ &= \int_{t_1\geq p} \boldsymbol{\phi}_1(t_1)\int_{t_2\leq t_1} f(t_1,t_2)\ \mathrm{d}t_2\ \mathrm{d}t_1. \end{aligned}$$

Since the above equation must hold for all p, by differentiating both sides with respect to p we get

$$\boldsymbol{\phi}_1(p) = -\frac{R'(p)}{\int_{t_2\leq p} f(p,t_2)\ \mathrm{d}t_2}.$$

Notice that

$$\begin{aligned} R'(p) &= \frac{d}{dp}(p\int_{\mathbf{t}\in T(p)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}) \\ &= \int_{\mathbf{t}\in T(p)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} + p\frac{d}{dp}(\int_{\mathbf{t}\in T(p)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}) \\ &= \int_{\mathbf{t}\in T(p)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} + p\frac{d}{dp}(\int_{t_1\geq p}\int_{t_2\leq t_1} f(t_1,t_2)\ \mathrm{d}t_2\ \mathrm{d}t_1) \\ &= \int_{\mathbf{t}\in T(p)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} - p\int_{t_2\leq p} f(p,t_2)\ \mathrm{d}t_2, \end{aligned}$$

and as a result,

$$R'(1) = -\int_{t_2\leq 1} f(1,t_2)dt_2.$$

Plugging in the definition of $\boldsymbol{\phi}_1$ implies that

$$\boldsymbol{\phi}_1(1) = 1,$$

and therefore $\boldsymbol{\alpha}_1(1,t_2) = f(1,t_2)(1-\boldsymbol{\phi}_1(1,t_2)) = 0$, as required by one of the boundary conditions of Lemma 2.

Having found a closed form for $\boldsymbol{\phi}_1$, we define $\boldsymbol{\phi}_2$ such that the conditions that $\boldsymbol{\alpha}\cdot\eta = 0$ when $t_2 = 0$, and also that $\nabla\cdot\boldsymbol{\alpha} = -f$ everywhere, are automatically satisfied,

$$\boldsymbol{\phi}_2(t_1,t_2)f(t_1,t_2) \quad = \quad t_2 f(t_1,t_2) + \int_{y=0}^{t_2} f(t_1,y) + \frac{d}{dt_1}[(t_1-\boldsymbol{\phi}_1(t_1,y))f(t_1,y)]\ \mathrm{d}y.$$

We intentionally do not simplify the definition of $\boldsymbol{\phi}_2$, since as we will see later it is irrelevant in the ex-post optimization, which we do by showing that it is in a sense dominated by $\boldsymbol{\phi}_1$. So we only verify that $\boldsymbol{\alpha}$ satisfies all the conditions of Lemma 2. Notice that given the above definition we have,

$$\frac{d}{dt_2}[(t_2 - \boldsymbol{\phi}_2(t_1,t_2))f(t_1,t_2)] = -f(t_1,t_2) - \frac{d}{dt_1}[(t_1 - \boldsymbol{\phi}(t_1,t_2))f(t_1,t_2)],$$

and as a result,

$$\begin{aligned} \nabla \cdot \boldsymbol{\alpha} &= \frac{d}{dt_1}[(t_1 - \boldsymbol{\phi}(t_1,t_2))f(t_1,t_2)] + \frac{d}{dt_2}[(t_2 - \boldsymbol{\phi}_2(t_1,t_2))f(t_1,t_2)] \\ &= -f(t), \end{aligned}$$

as required. This completes the proof that $\boldsymbol{\alpha}$ satisfies all the requirements of Lemma 2.

In order to apply Theorem 5, we need to show that for some p, the allocation function $\mathbf{x}_p$ optimizes $\boldsymbol{\phi}$ ex-post. Ex-post optimization of $\mathbf{x} \cdot \boldsymbol{\phi}$ will result in $\mathbf{x} = (1,0)$ whenever $\boldsymbol{\phi}_1 \geq 0$ and $\boldsymbol{\phi}_1 \geq \boldsymbol{\phi}_2$, and will result in $\mathbf{x} = (0,0)$ whenever $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2 \leq 0$. Since $\boldsymbol{\phi}_1(\mathbf{t})$ is a function of t_1 only, as long as $\boldsymbol{\phi}_1(t_1)$ is a monotone non-decreasing function of t_1, there will exist p^* such that $\boldsymbol{\phi}_1(t_1) \geq 0$ whenever $t_1 \geq p^*$, and $\boldsymbol{\phi}_1(t_1) \leq 0$ otherwise. So assuming monotonicity, we only need to verify that $\boldsymbol{\phi}_1(\mathbf{t}) \geq \boldsymbol{\phi}_2(\mathbf{t})$ whenever $t_1 \geq p^*$, and $\boldsymbol{\phi}_2(\mathbf{t}) \leq 0$ whenever $t_1 \leq p^*$. In order to do so, we will specify a condition on the distribution under which the following property holds.

$$\boldsymbol{\phi}_2(t_1,t_2) \leq \frac{t_2}{t_1}\boldsymbol{\phi}_1(t_1,t_2). \tag{2}$$

To specify the condition on distributions, it is useful to view the distribution in the *max-ratio* (MR) coordinates. In particular the max-ratio representation $\bar{f}$ of a distribution f is defined to be

$$\bar{f}(t_1,\theta) = f(t_1,\theta t_1).$$

Equivalantly,

$$f(t_1,t_2) = \bar{f}(t_1,t_2/t_1).$$

We call a distribution MR-log-supermodular, if its max-ratio representation is log-supermodular,

$$\bar{f}(t_1,\theta_1) \times \bar{f}(t_1',\theta_1') \leq \bar{f}(t_1,\theta_1') \times \bar{f}(t_1',\theta_1), \qquad \forall t_1 \leq t_1', \theta_1 \geq \theta_1'.$$

Notice that, for example, *MR-independent* distributions are MR-log-supermodular. A MR-independent distribution $\bar{f}$ is such that $\bar{f}(t_1,\theta_1) = f_1(t_1) \times f_2(\theta_1)$ (for arbitrary f_1 and f_2), and is MR-log-supermodular because

$$\begin{aligned} \bar{f}(t_1,\theta_1) \times \bar{f}(t_1',\theta_1') &= f_1(t_1)\, f_2(\theta_1)\, f_1(t_1')\, f_2(\theta_1') \\ &= f_1(t_1)\, f_2(\theta_1')\, f_1(t_1')\, f_2(\theta_1) \\ &= \bar{f}(t_1,\theta_1') \times \bar{f}(t_1',\theta_1). \end{aligned}$$

We start proving that Inequality (2) holds by observing a property of MR-log-supermodular functions.

Lemma 6. *If f is a MR-log-supermodular function, then for any t_1 and θ,*

$$\frac{d}{dt_1}\left(\frac{\int_{y=0}^{t_1\theta} f(t_1,y)\ \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1,y)\ \mathrm{d}y}\right) \leq 0,$$

with equality if distribution is MR-independent.

Proof. To give some intuition about why the above property holds, we first verify it for MR-independent distributions. The complete proof is given in Appendix B. For such a distribution we have

$$\begin{aligned}
\frac{\int_{y=0}^{t_1\theta} f(t_1,y)\ \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1,y)\ \mathrm{d}y} &= \frac{\int_{z=0}^{\theta} \bar{f}(t_1,z)t_1\ \mathrm{d}z}{\int_{z=0}^{1} \bar{f}(t_1,z)t_1\ \mathrm{d}z} \\
&= \frac{\int_{z=0}^{\theta} f_1(t_1)f_2(z)t_1\ \mathrm{d}z}{\int_{z=0}^{1} f_1(t_1)f_2(z)t_1\ \mathrm{d}z} \\
&= \frac{\int_{z=0}^{\theta} f_2(z)\ \mathrm{d}z}{\int_{z=0}^{1} f_2(z)\ \mathrm{d}z}.
\end{aligned}$$

Notice that above is not a function of t_1, and therefore we have shown that for MR-independent distributions,

$$\frac{d}{dt_1}\left(\frac{\int_{y=0}^{t_1\theta} f(t_1,y)\ \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1,y)\ \mathrm{d}y}\right) = 0. \tag{3}$$

□

We are now ready to prove Inequality (2). The proof is presented in Appendix B.

Lemma 7. *If a distribution is MR-log-supermodular, then $\boldsymbol{\phi}_2(t_1,t_2) \leq \frac{t_2}{t_1}\boldsymbol{\phi}_1(t_1,t_2)$, with equality if the distribution is MR-independent.*

To understand MR-log-supermodular distributions better, we next show properties under which product distributions are MR-log-supermodular.

We call a function f geometric-geometric (GG) convex if

$$\forall 0 \leq \lambda \leq 1, v_1, v_2,\ f(v_1^{\lambda}v_2^{1-\lambda}) \geq f(v_1)^{\lambda}f(v_2)^{1-\lambda}.$$

For example, the function $f(x) = x^k$ is GG-convex. More generally, for any convex function f and constant c, the function $c \cdot e^{f(\log(x))}$ is GG-convex.

Lemma 8. *A symmetric product distribution is MR-log-supermodular if it is GG-convex*

Theorem 9. *Uniform pricing is the optimal auction if the distribution is symmetric and MR-log-supermodular, and the revenue function $R(p)$ is concave.*

As stated before, we can improve the above theorem by using an ironed virtual value function and applying Theorem 4.

Theorem 10. *Uniform pricing is the optimal auction if the distribution is symmetric and MR-log-supermodular.*

In Appendix A we show that the above result can be generalized to $m \geq 2$ items, using the natural generalization of MR-log-supermodular functions to $m \geq 2$ dimensions. The max-ratio representation of the density function f is a function $\bar{f}$ satisfying

$$\bar{f}(t_1, \theta_2, \ldots, \theta_m) = f(t_1, \theta_2 t_1, \ldots, \theta_m t_1).$$

A density function is MR-log-supermodular if it's max-ratio representation is log-supermodular.

4.1 Bundle Pricing for General Distributions

We can use the generality of unit-demand settings and the above result to find conditions under which bundle pricing is optimal for general distributions. More precisely, notice that any problem can be converted to a unit-demand problem with dot-product distribution by letting $\mathbf{t}$ and $\mathbf{x}$ be vectors of size equal to the size of the outcome space, with $\mathbf{t}_i$ being the value for outcome i. For milt-item settings, the size of the outcome space is equal to 2^m. The input to the problem is a density function $f(t_1, \ldots, t_k)$ over vectors of size $k = 2^m$. We assume that the density function is *monotone*, meaning that the density of a type is non-zero only if the valuation for the grand bundle is at least as much as the value for any other subset of items. The max-ratio representation of the density function is $\bar{f}(t_1, \theta_2, \ldots, \theta_k)$, where t_1, normalized to be at most 1, is the value for the grand bundle, and $\theta_i \leq 1$ is the ratio of the value for the i-th subset, over the value of the grand bundle. Now a bundle pricing mechanism corresponds to a mechanism in the unit demand setting which only sells the favorite outcome, which is the grand bundle. We therefore have the following theorem.

Theorem 11. *Bundle pricing is optimal if the density function f is MR-log-supermodular.*

5 Optimality of Bundle Pricing for Additive Preferences

In this section we provide sufficient and necessary conditions for optimality of bundle pricing in additive settings. It is easiest to express the results of this section when the set of types is normalized to be contained in $T = \{(t_1, t_2) | t_1, t_2 \geq 0, t_1 + t_2 \leq 1\}$. Similar to Section 4, by symmetry we only focus on constructing a proof for the subset of type $\mathbf{t} \in T$ such that $t_1 \geq t_2$.

Recall that we in order to apply Theorem 4, we need to construct $\boldsymbol{\alpha}$ and $\bar{\boldsymbol{\phi}}$ satisfying its 3 required properties. In order to satisfy the third property, we construct $\bar{\boldsymbol{\phi}}$ such that the expected virtual value of $\mathbf{x}_p$, the allocation function of posting a price p for the bundle, is equal to the revenue of posting a price p for the bundle, $R(p) = p \int_{\mathbf{dt} \in T(p)} f(\mathbf{t}) \, \mathrm{d}\mathbf{t}$, where $T(p) = \{(t_1, t_2) | (t_1, t_2) \in T, t_1 + t_2 \geq p\}$.

$$\begin{aligned} R(p) &= \int_{\mathbf{t} \in T} (\mathbf{x}_p \cdot \bar{\boldsymbol{\phi}} f)(\mathbf{t}) \, \mathrm{d}\mathbf{t} \\ &= p \int_{\mathbf{t} \in T(p)} [\bar{\boldsymbol{\phi}}_1 + \bar{\boldsymbol{\phi}}_2](\mathbf{t}) f(\mathbf{t}) \, \mathrm{d}\mathbf{t} \end{aligned}$$

The second equality followed because $\mathbf{x}(\mathbf{t}) = (1,1)$ for $\mathbf{t} \in T(p)$, and $\mathbf{x}(\mathbf{t}) = (0,0)$ otherwise. In parallel to section Section 4, we further assume that $(\bar{\boldsymbol{\phi}}_1 + \bar{\boldsymbol{\phi}}_2)(\mathbf{t}) = \lambda(t_1 + t_2)$, for some function λ. As a result, by performing a change of variables $s = t_1 + t_2$, and $\theta = t_2/t_1$, we have

$$\begin{aligned} R(p) &= \int_{s \geq p}^{1} \int_{\theta=0}^{1} \lambda(s) \bar{f}(s,\theta) \frac{s}{1+\theta} \, \mathrm{d}\theta \, \mathrm{d}s \\ &= \int_{s \geq p}^{1} s\lambda(s) \int_{\theta=0}^{1} \frac{\bar{f}(s,\theta)}{1+\theta} \, \mathrm{d}s. \end{aligned}$$

We require above to hold for all p, and therefore we differentiate the above equation with respect to p to specify λ, and therefore $\bar{\boldsymbol{\phi}}$.

$$\lambda(p) = -\frac{R'(p)}{p \int_{\theta=0}^{1} \frac{\bar{f}(p,\theta)}{1+\theta} \, \mathrm{d}\theta}.$$

We further assume that $\bar{\boldsymbol{\phi}}_2 = \frac{t_2}{t_1} \bar{\boldsymbol{\phi}}_1$. This will define $\bar{\boldsymbol{\phi}}_1$ and $\bar{\boldsymbol{\phi}}_2$ as follows

$$\begin{aligned} \bar{\boldsymbol{\phi}}_1(\mathbf{t}) &= -\frac{t_1}{t_1 + t_2} \frac{R'(p)}{p \int_{\theta=0}^{1} \frac{\bar{f}(p,\theta)}{1+\theta} \, \mathrm{d}\theta} \\ \bar{\boldsymbol{\phi}}_2(\mathbf{t}) &= -\frac{t_2}{t_1 + t_2} \frac{R'(p)}{p \int_{\theta=0}^{1} \frac{\bar{f}(p,\theta)}{1+\theta} \, \mathrm{d}\theta} \end{aligned}$$

Notice that $\bar{\boldsymbol{\phi}}_1$ and $\bar{\boldsymbol{\phi}}_2$ have the same sign as λ, and therefore as long as λ is a non-decreasing function of p, there exists a price p^* such that the allocation corresponding to price p^* ex-post optimizes $\bar{\boldsymbol{\phi}}$. So in order to complete the construction of proof we should construct $\boldsymbol{\alpha}$ such that $\bar{\boldsymbol{\phi}}$ and $\boldsymbol{\alpha}$ satisfy property 1 of Theorem 4. That is, to show that for all incentive compatible $\mathbf{x}$, we have

$$\int_{\mathbf{t}} \mathbf{x}(\mathbf{t})(\mathbf{t}f - \boldsymbol{\alpha}(\mathbf{t}) \, \mathrm{d}\mathbf{t} \leq \int_{\mathbf{t}} \mathbf{x}(\mathbf{t}) \cdot \bar{\boldsymbol{\phi}}(\mathbf{t}) f(\mathbf{t}) \, \mathrm{d}\mathbf{t}.$$

To simplify exposition we will assume $\mathbf{t}f - \boldsymbol{\alpha}(\mathbf{t}) = \boldsymbol{\phi}(\mathbf{t}) f(\mathbf{t})$, after which our goal becomes to show that the following expression is at most zero, for all incentive compatible $\mathbf{x}$,

$$\int_{\mathbf{t}} \mathbf{x}(\mathbf{t}) \cdot [\boldsymbol{\phi}(\mathbf{t}) - \bar{\boldsymbol{\phi}}(\mathbf{t})] f(\mathbf{t}) \, \mathrm{d}\mathbf{t}.$$

We will now perform a change of variables $s = t_1 + t_2$, and $\theta = t_2/t_1$,

$$\int_{s=0}^{1} \int_{\theta=0}^{1} \mathbf{x}(s,\theta) \cdot [\boldsymbol{\phi} - \bar{\boldsymbol{\phi}}](s,\theta) f(s,\theta) \frac{s}{1+\theta} \, \mathrm{d}s \, \mathrm{d}\theta.$$

We can integrate by parts and write the above expression as

$$\int_{s=0}^{1}\int_{\theta=0}^{1}\mathbf{x}(s,\theta)\cdot[\boldsymbol{\phi}-\bar{\boldsymbol{\phi}}](s,\theta)f(s,\theta)\frac{s}{1+\theta}\,\mathrm{d}s\,\mathrm{d}\theta=\int_{s=0}^{1}\int_{\theta=0}^{1}\mathbf{x}(s,\theta)\cdot\frac{d}{d\theta}\int_{y=\theta}^{1}[\boldsymbol{\phi}-\bar{\boldsymbol{\phi}}](s,y)f(s,y)\frac{s}{1+y}\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}\theta$$
$$=\int_{s=0}^{1}\int_{\theta=0}^{1}\frac{d}{d\theta}\mathbf{x}(s,\theta)\cdot\int_{y=\theta}^{1}[\boldsymbol{\phi}-\bar{\boldsymbol{\phi}}](s,y)f(s,y)\frac{s}{1+y}\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}\theta$$

Define

$$\gamma(s,\theta)=\int_{y=\theta}^{1}[\boldsymbol{\phi}-\bar{\boldsymbol{\phi}}](s,y)f(s,y)\frac{s}{1+y}\,\mathrm{d}y,$$

which allows us to write

$$\int_{s=0}^{1}\int_{\theta=0}^{1}\mathbf{x}(s,\theta)\cdot[\boldsymbol{\phi}-\bar{\boldsymbol{\phi}}](s,\theta)f(s,\theta)\frac{s}{1+\theta}\,\mathrm{d}s\,\mathrm{d}\theta=\int_{s=1}^{1}\int_{\theta=0}^{1}\frac{d}{d\theta}\mathbf{x}(s,\theta)\cdot\gamma(s,\theta)\,\mathrm{d}s\,\mathrm{d}\theta.$$

The following two lemmas show that the above expression is negative for all incentive compatible allocation under our regularity assumption, completing the proof.

Lemma 12. *The allocation of any incentive compatible mechanism satisfies*

$$\frac{d}{d\theta}\boldsymbol{x}(s,\theta)\cdot(1,-1)\leq 0$$

Proof. Incentive compatibility implies that for any s,θ, and ϵ,

$$(\mathbf{x}(s,\theta+\epsilon)-\mathbf{x}(s,\theta))\cdot(-1,1)\frac{\theta\epsilon}{s}\geq 0.$$

Letting ϵ approach zero implies the claim. □

We say that a distribution is *sum-ratio* log-submodular (SR-log-submodular) if its sum-ratio representation is log-supermodular, that is,

$$\bar{f}(s,\theta)\times\bar{f}(s',\theta')\geq\bar{f}(s',\theta)\times\bar{f}(s,\theta'),\qquad\forall s\leq s',\theta\leq\theta'$$

Lemma 13. *If the distribution is SR-log-submodular, then there exist $\boldsymbol{\alpha}$ satisfying conditions of Lemma 2 such that $\gamma_1=-\gamma_2$, where $\gamma_1\geq 0$.*

Proof. We will show that the condition that $\gamma_1=-\gamma_2$, together with conditions of Lemma 2 uniquely specify $\boldsymbol{\alpha}$ (see Appendix). □

Theorem 14. *Bundle pricing is optimal for any SR-log-submodular distribution.*

6 Multi Agent Extensions

6.1 The Setting

A multi-agent problem is defined by n agents, each agent κ associated with a distribution f^κ, and a feasibility setting $\mathcal{S} \subseteq \{0,1\}^{n\times m}$. Types of agents are drawn independent of each other from corresponding distributions. A multi-agent mechanism is a sequence of allocation functions $(\bar{\mathbf{x}}^1(\mathbf{t}^1,\ldots,\mathbf{t}^n),\ldots,\bar{\mathbf{x}}^n(\mathbf{t}^1,\ldots,\mathbf{t}^m)) \in \mathcal{S}$ in which $\bar{\mathbf{x}}^\kappa(\mathbf{t}^1,\ldots,\mathbf{t}^n) \in \mathbb{R}^m$ is the allocation of agent κ, together with a sequence of payment function $(p^1(\mathbf{t}^1,\ldots,\mathbf{t}^n),\ldots,p^n(\mathbf{t}^1,\ldots,\mathbf{t}^n)) \in \mathbb{R}^n$ in which $p^\kappa(\mathbf{t}^1,\ldots,\mathbf{t}^n) \in \mathbb{R}$ is the payment of agent κ. A multi-agent mechanism is *ex-post individually rational* (EPIR) if the utility of every type of every agent is always positive,

$$\mathbf{t}^\kappa \cdot \bar{\mathbf{x}}^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa}) - p^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa}) \geq 0, \qquad \forall \kappa, \mathbf{t}^\kappa, \mathbf{t}^{-\kappa},$$

where $\mathbf{t}^{-\kappa}$ is the vector of other agents types. A multi-agent mechanism is *dominant strategy incentive compatible* (DSIC) if no type of any agent increases his utility by misreporting,

$$\mathbf{t}^\kappa \cdot \bar{\mathbf{x}}^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa}) - p^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa}) \geq \mathbf{t}^\kappa \cdot \bar{\mathbf{x}}^\kappa(\hat{\mathbf{t}}^\kappa,\mathbf{t}^{-\kappa}) - p^\kappa(\hat{\mathbf{t}}^\kappa,\mathbf{t}^{-\kappa}), \qquad \forall \kappa, \mathbf{t}^\kappa, \hat{\mathbf{t}}^\kappa, \mathbf{t}^{-\kappa}.$$

The *interim mechanism* for agent κ is a single-agent mechanism $(\mathbf{x}^\kappa, p^\kappa)$ defined to be

$$\begin{aligned} \mathbf{x}^\kappa(\mathbf{t}^\kappa) &= E_{\mathbf{t}^{-\kappa}\sim f^{-\kappa}}[\bar{\mathbf{x}}^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa})], \\ p^\kappa(\mathbf{t}^\kappa) &= E_{\mathbf{t}^{-\kappa}\sim f^{-\kappa}}[p^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa})], \end{aligned}$$

where $\mathbf{t}^{-\kappa}$ is the vector of other agents types, drawn at random from $f^{-\kappa} := f^1 \times \ldots \times f^{\kappa-1} \times f^{\kappa+1} \times \ldots \times f^n$. A mechanism is *interim individually rational* (IIR) if the interim mechanism of every agent is individually rational. A mechanism is *Bayesian incentive compatible* (BIC) if the interim mechanism of every agent is incentive compatible. Notice that EPIR implies IIR, and DSIC implies BIC. The goal is to design feasible IIC and IIR mechanisms that maximize expected payments. We however prove stronger results that the optimal mechanisms are in fact EPIR and DSIC.

6.2 Problem Formulation

A single agent mechanism $(\mathbf{x}, p)$ defines a utility function $u(\mathbf{t}) = \mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - p(\mathbf{t})$. The following lemma connects the utility function of an IC mechanism with its allocation function.

Lemma 15 (Manelli and Vincent (2006))**.** *Function u is the utility function of an agent in an incentive compatible mechanism if and only if u is convex, and in that case, the agent's allocation is $\mathbf{x}(\mathbf{t}) = \nabla u(\mathbf{t})$.*

Notice that the payment function can be defined using the utility function and the allocation function as $p(\mathbf{t}) = \mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - u(\mathbf{t})$. Using the above lemma, we can write payment to be $p(\mathbf{t}) = \mathbf{t} \cdot \nabla u(\mathbf{t}) - u(\mathbf{t})$. This allows us to write the following optimization problem as our revenue maximization problem, which is the starting point of the analysis of this paper.

$$\max_{\bar{\mathbf{x}}^1,\ldots,\bar{\mathbf{x}}^n,u^1,\ldots,u^n} \quad \sum_\kappa \int_{\mathbf{t}} [\mathbf{t}\cdot\nabla u^\kappa(\mathbf{t}) - u^\kappa(\mathbf{t})] f^\kappa(\mathbf{t})\,\mathrm{d}\mathbf{t} \tag{4}$$
$$\forall \kappa, \mathbf{x}^\kappa = \nabla u^\kappa; u^\kappa \text{ is convex},$$
$$\forall \mathbf{t}, (\bar{\mathbf{x}}^1(\mathbf{t}^1,\ldots,\mathbf{t}^n),\ldots,\bar{\mathbf{x}}^n(\mathbf{t}^1,\ldots,\mathbf{t}^n)) \text{ is feasible allocation.}$$

Notice that when $m = 1$, the above program is equivalent to the following familiar form of

$$\max_{\bar{\mathbf{x}}^1,\ldots,\bar{\mathbf{x}}^n} \quad \sum_i \int_v [v\mathbf{x}^\kappa(v) - \int_{z\le v} \mathbf{x}^\kappa(z)dz] f(v)\,\mathrm{d}v$$
$$\forall \kappa, \mathbf{x}^\kappa \text{ is monotone non-decreasing},$$
$$\forall \mathbf{t}, (\bar{\mathbf{x}}^1(\mathbf{t}^1,\ldots,\mathbf{t}^n),\ldots,\bar{\mathbf{x}}^n(\mathbf{t}^1,\ldots,\mathbf{t}^n)) \text{ is feasible allocation.}$$

Now assume that for each agent κ, we have constructed $\boldsymbol{\alpha}^\kappa(\mathbf{t})$ according to Lemma 2. The above analysis and Program 4 tells us that we can write revenue as

$$\begin{aligned}
\sum_\kappa \int_{\mathbf{t}^\kappa} \mathbf{x}^\kappa(\mathbf{t}^\kappa)\cdot\boldsymbol{\phi}^\kappa(\mathbf{t}^\kappa) f^\kappa(\mathbf{t}^\kappa)\,\mathrm{d}\mathbf{t}^\kappa &= \sum_\kappa \int_{\mathbf{t}^\kappa} \Big[\int_{\mathbf{t}^{-\kappa}} \bar{\mathbf{x}}^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa}) f^{-\kappa}(\mathbf{t}^{-\kappa})\,\mathrm{d}\mathbf{t}^{-\kappa}\Big]\cdot\boldsymbol{\phi}^\kappa(\mathbf{t}^\kappa) f^\kappa(\mathbf{t}^\kappa)\,\mathrm{d}\mathbf{t}^\kappa \\
&= \sum_\kappa \int_{\mathbf{t}} \bar{\mathbf{x}}^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa})\cdot\boldsymbol{\phi}^\kappa(\mathbf{t}^\kappa) f(\mathbf{t})\,\mathrm{d}\mathbf{t} \\
&= \int_{\mathbf{t}} \Big[\sum_\kappa \bar{\mathbf{x}}^\kappa(\mathbf{t}^\kappa,\mathbf{t}^{-\kappa})\cdot\boldsymbol{\phi}^\kappa(\mathbf{t}^\kappa)\Big] f(\mathbf{t})\,\mathrm{d}\mathbf{t}.
\end{aligned}$$

We say that a mechanism $\bar{\mathbf{x}}$ optimizes an objective $\boldsymbol{\phi}$ *ex-post*, if for any input $\mathbf{t}^1,\ldots,\mathbf{t}^m$ it selects an outcome that maximizes $\sum_\kappa \bar{\mathbf{x}}^\kappa(\mathbf{t}^1,\ldots,\mathbf{t}^m)\cdot\boldsymbol{\phi}^\kappa(\mathbf{t}^\kappa)$.

Notice that in a multi-agent problem, the existence of $\boldsymbol{\alpha}^\kappa$ depends only on the distribution f^κ. As a result, in the rest of this section we study construction of $\boldsymbol{\alpha}^\kappa$ for single agents, and drop the index κ. By showing the existence of proofs for each agent separately, we can apply the meta theorem to get the structure of the multi-agent optimal solution.

6.3 Ex-post Optimization and Ironing

Having proved that $\boldsymbol{\alpha}$ is a valid amortization of revenue by proving it satisfies conditions of Lemma 2, we now study allocation rules that optimize (α_1,α_2) ex-post. First we define *service-constrained* settings. A service-constrained environment is characterized by a set system $\mathcal{S}\subseteq 2^n$. An allocation is feasible if $\{i|\sum_{j=m} x_{ij} = 1\}\in\mathcal{S}$.

Notice that $\alpha_2 \le \frac{t_2}{t_1}\alpha_1$ implies that when $\alpha_1 \ge 0$, we have $\alpha_2 \le \alpha_1$, and when $\alpha_1 \le 0$, we also have $\alpha_2 \le 0$. This implies that in service-constrained settings, an ex-post optimization of $\boldsymbol{\alpha}$ implies that the allocation can be positive only for the favorite item. In addition, if $\alpha(t)$ is a non-decreasing function of t_1, then so will be the allocation probability for the favorite item. This implies that the allocation rule is incentive compatible because

$$\begin{aligned}
(\mathbf{x}(\mathbf{t}) - \mathbf{x}(\mathbf{t}'))\cdot(\mathbf{t}-\mathbf{t}') &= (\mathbf{x}_1(\mathbf{t}) - \mathbf{x}_1(\mathbf{t}'))(t_1 - t_1') \\
&\ge 0,
\end{aligned}$$

which shows the convexity of the utility function, which in turn proves incentive compatibility (see Lemma 15). We have therefore proved the following lemma.

Lemma 16. *If $\alpha(\boldsymbol{t})$ is an increasing function of t_1, and $\alpha_1 \geq \alpha_2$ whenever $\alpha_2 \geq 0$, ex-post optimization of (α_1, α_2) in any service-constrained settings results in an incentive compatible allocation.*

Unfortunately, as it happens in single-dimensional settings, the derivative of revenue is not necessarily monotone. However, we can fix this problem by *ironing* the virtual value function in a similar manner to the single-dimensional problem. An ironing of multi-dimensional virtual values should be performed carefully in order to keep the properties of virtual values. More precisely, we would like to maintain the property that $\alpha_2 \leq \alpha_1$ when $\alpha_2 \geq 0$. In particular, we can prove the following lemma (proof is in Appendix B).

Lemma 17. *There exists an ironed virtual value function $\hat{\boldsymbol{\phi}}$ such that $\mathbf{x}\cdot(\mathbf{t}f(\mathbf{t})-\hat{\boldsymbol{\phi}}) \geq \mathbf{x}\cdot(\mathbf{t}f(\mathbf{t})-\boldsymbol{\alpha})$ for any incentive compatible allocation. In addition, the allocation rule $\hat{\mathbf{x}}$ that ex-post optimizes $\hat{\boldsymbol{\phi}}$ satisfies $\hat{\mathbf{x}} \cdot (\mathbf{t}f(\mathbf{t}) - \hat{\boldsymbol{\phi}}) = \hat{\mathbf{x}} \cdot (\mathbf{t}f(\mathbf{t}) - \boldsymbol{\alpha})$.*

The above lemma, together with Lemma 16 implies our main theorem.

Theorem 18. *Ex-post optimization of the ironed virtual value function is optimal for MR-log-supermodular distributions in service-constrained settings. In the optimal allocation the probability of receiving an item is non-zero only if the item is the favorite item.*

To show the strength of the above result, we complement it with an almost matching converse result. In particular, we prove that for $\sqrt{2}$MR-log-submodular functions, uniform pricing is not optimal even for a single agent. We start by defining $\sqrt{2}$MR-log-submodular distributions. A distribution is $\sqrt{2}$MR-log-submodular if the max-ratio representation of it is $\sqrt{2}$-log-submodular.

$$\bar{f}(t_1, \theta_1) \times \bar{f}(t_1', \theta_1') > \sqrt{2} \times \bar{f}(t_1, \theta_1') \times \bar{f}(t_1', \theta_1), \qquad \forall t_1 \leq t_1', \theta_1 \geq \theta_1'.$$

Theorem 19. *Uniform prices are not optimal for $\sqrt{2}$MR-log-submodular distributions.*

Proof. Assume for contradiction that a price p^* is optimal for such a distribution. Such a price must satisfy the local condition of optimality, that is,

$$\int_{T(p^*,1)} f(\mathbf{t}) \, \mathrm{d}\mathbf{t} = p^* \int_{C(p^*,1)} f(\mathbf{t}) \, \mathrm{d}\mathbf{t}.$$

In addition, adding a lottery with price $p^* - \epsilon$ that sells each item with probability $1/2$ should not increase revenue for any ϵ. In the limit, this implies that

$$p^* f(p^*, p^*) \leq \sqrt{2} \int_{x=p^*}^{1} f(x, x) \, \mathrm{d}x.$$

We will show that given strict $\sqrt{2}$MR-log-submodularity both of the above conditions can not be simultaneously satisfied. By multiplying the conditions we get

$$f(p^*, p^*) \int_{T(p^*,1)} f(\mathbf{t}) \, \mathrm{d}\mathbf{t} \leq \sqrt{2} \int_{C(p^*,1)} f(\mathbf{t}) \, \mathrm{d}\mathbf{t} \times \int_{x=p^*}^{1} f(x, x) \, \mathrm{d}x.$$

But notice that given strict $\sqrt{2}$MR-log-submodularity

$$\begin{aligned}\sqrt{2}\int_{C(p^*,1)} f(\mathbf{t})\,\mathrm{d}\mathbf{t} \times \int_{x=p^*}^{1} f(x,x)\,\mathrm{d}x &= \sqrt{2}\int_{y=0}^{p^*} f(p^*,y)\,\mathrm{d}y \times \int_{x=p^*}^{1} f(x,x)\,\mathrm{d}x \\ &= \sqrt{2}\int_{x=p^*}^{1}\int_{y=0}^{1} \bar{f}(p^*,y)\bar{f}(x,x)y\,\mathrm{d}y\,\mathrm{d}x \\ &< \int_{x=p^*}^{1}\int_{y=0}^{1} \bar{f}(p^*,p^*)\bar{f}(x,y)y\,\mathrm{d}y\,\mathrm{d}x \\ &= f(p^*,p^*)\int_{T(p^*,1)} f(\mathbf{t})\,\mathrm{d}\mathbf{t},\end{aligned}$$

which is a contradiction. □

References

Alaei, S., Fu, H., Haghpanah, N., and Hartline, J. (2013). The simple economics of approximately optimal auctions. In *FOCS*.

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica*, 64(1):51–75.

Armstrong, M. (1999). Price discrimination by a many-product firm. *The Review of Economic Studies*, 66(1):151–168.

Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2013). Mechanism design via optimal transport. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 269–286. ACM.

Hart, S. and Nisan, N. (2013). The menu-size complexity of auctions. *arXiv preprint arXiv:1304.6116*.

Hartline, J. D. (2012). Bayesian mechanism design. *Foundations and Trends® in Theoretical Computer Science*, 8(3):143–263.

Manelli, A. M. and Vincent, D. R. (2006). Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1–35.

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):pp. 58–73.

Pavlov, G. (2011). A property of solutions to linear monopoly problems. *The BE Journal of Theoretical Economics*, 11(1).

Rochet, J. (1985). The taxation principle and multi-time hamilton-jacobi equations. *Journal of Mathematical Economics*, 14(2):113 – 128.

Thanassoulis, J. (2004). Haggling over substitutes. *J. Economic Theory*, 117(2):217–245.

Wang, Z. and Tang, P. (2013). Optimal mechanisms with simple menus. *arXiv preprint arXiv:1311.5966*.

A Extension for $m > 2$ Items

In this section, we extend the definition of $\boldsymbol{\alpha}$ given in Section 4 from two items to $m > 2$. The approach to derive such general $\boldsymbol{\alpha}$ is similar to that of Section 4, and therefore in this section we only define $\boldsymbol{\alpha}$ and verify its properties. Assume that we are in a partition of the space in which $t_1 \geq t_i$ for all i. Define $T(t_1, \theta_2, \ldots, \theta_m) = \{\mathbf{t}' | t'_1 \geq t_1, t'_i \leq \theta_i t_1, \forall i \geq 2\}$. Define $C(t_1, \theta_2, \ldots, \theta_m) = \{\mathbf{t}' | t'_1 = t_1, t'_i \leq t'_i \theta_i, \forall i \geq 2\}$. Now define

$$\alpha_1(\mathbf{t}) = t_1 - \frac{\int_{\mathbf{t}' \in T(t_1,1,\ldots,1)} f(\mathbf{t}') \, \mathrm{d}\mathbf{t}'}{\int_{\mathbf{t}' \in C(t_1,1,\ldots,1)} f(\mathbf{t}') \, \mathrm{d}\mathbf{t}'}$$

We also define α_i for $i \geq 2$ as follows

$$\alpha_i(\mathbf{t}) f(\mathbf{t}) = t_i f(\mathbf{t}) + \frac{1}{m-1} \int_{y=0}^{t_i} f(y, \mathbf{t}_{-i}) + \frac{d}{dt_1}[(t_1 - \alpha_1(y, \mathbf{t}_{-i})) f(y, \mathbf{t}_{-i})] \, \mathrm{d}y$$

The above definition implies that

$$\frac{d}{dt_i}\Big(t_i f(\mathbf{t}) - \alpha_i(\mathbf{t}) f(\mathbf{t})\Big) = -\frac{1}{m-1}\Big(f(\mathbf{t}) + \frac{d}{dt_1}[(t_1 - \alpha_1(\mathbf{t})) f(\mathbf{t})]\Big).$$

As a result,

$$\begin{aligned}
\nabla \cdot \phi &= \sum_i \frac{d}{dt_i}\Big(t_i f(\mathbf{t}) - \alpha_i(\mathbf{t}) f(\mathbf{t})\Big) \\
&= \frac{d}{dt_1}[(t_1 - \alpha_1(\mathbf{t})) f(\mathbf{t})] - (m-1) \times \frac{1}{m-1}\Big(f(\mathbf{t}) + \frac{d}{dt_1}[(t_1 - \alpha_1(\mathbf{t})) f(\mathbf{t})]\Big) \\
&= -f
\end{aligned}$$

We now verify that ϕ satisfies the boundary conditions. This holds because when $t_i = 0$, $\alpha_i(\mathbf{t}) = 0$, and also when $t_1 = 1$, $\alpha_1(\mathbf{t}) = f(\mathbf{t})$.

Finally, we verify that $\alpha_i(\mathbf{t}) \leq \frac{t_i}{t_1} \alpha_1(\mathbf{t})$. This is again done in a manner similar to Section 4. Fix values of $\theta_3, \ldots, \theta_m$, let $T_{\theta_3,\ldots,\theta_n}(t_1, \theta_2)$ be the projection of type space into set of types such that each type $\mathbf{t}'$ satisfies $t'_1 \geq t_1$, $t'_1 \leq t'_2 \theta$, and $t'_i = t_1 \theta_i$. Now we can invoke the divergence theorem to conclude that $\alpha_i(\mathbf{t}) \leq \frac{t_i}{t_1} \alpha_1(\mathbf{t})$ if

$$\frac{d}{dt_1}\Big(\frac{\int_{y=0}^{t_1 \theta_2} f(t_1, y, \theta_3, \ldots, \theta_n) \, \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1, y, \theta_3, \ldots, \theta_n) \, \mathrm{d}y}\Big) \leq 0.$$

Notice that fixing $\theta_3, \ldots, \theta_n$, the above property is exactly what was required in two dimensions, which we showed follows from MR-log-supermodularity. As a result, if the function is MR-log-supermodular in every pair of variables t_1, θ_i for $i \geq 2$, then we $\alpha_i(\mathbf{t}) \leq \frac{t_i}{t_1} \alpha_1(\mathbf{t})$. Notice that this property is implied by MR-log-supermodularity of the distribution in all its variables, and therefore is a less demanding condition. We have therefore proved the following lemma.

Lemma 20. *If the distribution is MR-log-supermodular in every pair of variables t_1, θ_i for $i \geq 2$, then the revenue of an allocation can be upper bounded by*

$$\int_{\mathbf{t}} \mathbf{x} \cdot (1, \frac{t_2}{t_1}, \ldots, \frac{t_n}{t_1}) \alpha(t_1) f(\mathbf{t}) \, \mathrm{d}\mathbf{t},$$

where

$$\alpha_1(\mathbf{t}) = t_1 - \frac{\int_{\mathbf{t}' \in T(t_1,1,\ldots,1)} f(\mathbf{t}') \, \mathrm{d}\mathbf{t}'}{\int_{\mathbf{t}' \in C(t_1,1,\ldots,1)} f(\mathbf{t}') \, \mathrm{d}\mathbf{t}'}.$$

B Missing Proofs

B.1 Proof from Section 4

Proof of Lemma 6. We now prove the lemma for general MR-log-supermodular functions. We prove that for any θ, t, and t' such that $t_1 < t_1'$,

$$\frac{\int_{y=0}^{t_1 \theta} f(t_1, y) \, \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1, y) \, \mathrm{d}y} \leq \frac{\int_{y=0}^{t_1' \theta} f(t_1', y) \, \mathrm{d}y}{\int_{y=0}^{t_1'} f(t_1', y) \, \mathrm{d}y}.$$

The proof first converts the above form into max-ratio coordinates, applies MR-log-supermodularity, and then transforms back to the standard form. Before applying MR-log-supermodularity, we break down the integral set into two set, and apply MR-log-supermodularity to only one of the integrals. More particularly, notice that

$$
\begin{aligned}
&\int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y \times \int_{y=0}^{t_1'} f(t_1', y)\ \mathrm{d}y \\
&= \int_{z_1=0}^{\theta} \bar{f}(t_1, z_1)\ t_1\ \mathrm{d}z_1 \times \int_{z_2=0}^{1} \bar{f}(t_1', z_2)\ t_1'\ \mathrm{d}z_2 && \text{(change of variables)} \\
&= \int_{z_1=0}^{\theta}\int_{z_2=0}^{\theta} \bar{f}(t_1, z_1)\ t_1\ \bar{f}(t_1', z_2)\ t_1'\ \mathrm{d}z_2\ \mathrm{d}z_1 \\
&\quad + \int_{z_1=0}^{\theta}\int_{z_2=\theta}^{1} \bar{f}(t_1, z_1)\ t_1 \bar{f}(t_1', z_2)\ t_1'\ \mathrm{d}z_2\ \mathrm{d}z_1 && \text{(separate double integral into two sets)} \\
&\leq \int_{z_1=0}^{\theta}\int_{z_2=0}^{\theta} \bar{f}(t_1, z_1)\ t_1\ \bar{f}(t_1', z_2)\ t_1'\ \mathrm{d}z_2\ \mathrm{d}z_1 \\
&\quad + \int_{z_1=0}^{\theta}\int_{z_2=\theta}^{1} \bar{f}(t_1, z_2)\ t_1\ \bar{f}(t_1', z_1)\ t_1'\ \mathrm{d}z_2\ \mathrm{d}z_1 && \text{(apply MR-log-supermodularity)} \\
&= \int_{z_1=0}^{\theta}\int_{z_2=0}^{\theta} \bar{f}(t_1, z_2)\ t_1\ \bar{f}(t_1', z_1)\ t_1'\ \mathrm{d}z_1\ \mathrm{d}z_2 && \text{(rename variables } z_1 \text{ and } z_2\text{)} \\
&\quad + \int_{z_1=0}^{\theta}\int_{z_2=\theta}^{1} \bar{f}(t_1, z_2)\ t_1\ \bar{f}(t_1', z_1)\ t_1'\ \mathrm{d}z_2\ \mathrm{d}z_1 \\
&= \int_{z_2=0}^{1} \bar{f}(t_1, z_2)\ t_1\ \mathrm{d}z_2 \int_{z_1=0}^{\theta} \bar{f}(t_2, z_1)\ t_1'\ \mathrm{d}z_1 && \text{(merge integrals)} \\
&= \int_{y=0}^{t_1} f(t_1, y)\ \mathrm{d}y \times \int_{y=0}^{t_1'\theta} f(t_1', d)\ \mathrm{d}y.
\end{aligned}
$$

□

Proof of Lemma 7. We prove this property using the divergence theorem. In particular, we invoke the divergence theorem for vector field $\boldsymbol{\alpha}$ over the set $T(t, \theta)$, for all t and θ,

$$
-\int_{T(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} = \int_{\partial T(t_1,\theta)} \boldsymbol{\alpha}(\mathbf{t}) \cdot \eta(\mathbf{t})\ \mathrm{d}\mathbf{t}.
$$

We now analyze the right hand of the above equation. Notice that the curves $C_1(t_1, 0), C_1(t_1, \theta), C_2(t_1, \theta)$, and $C_2(1, \theta)$ collectively define the boundary of $T(t, 1)$. Recall that the boundary conditions of $\boldsymbol{\alpha}$ imply that $\boldsymbol{\alpha} \cdot \eta = 0$ on $C_1(t, 0)$ and $C_2(1, \theta)$. The divergence condition therefore reduces to

$$
-\int_{T(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} = \int_{C_2(t_1,\theta)} \phi(\mathbf{t}) \cdot \eta(\mathbf{t})\ \mathrm{d}\mathbf{t} + \int_{C_1(t_1,\theta)} \phi(\mathbf{t}) \cdot \eta(\mathbf{t})\ \mathrm{d}\mathbf{t}.
$$

Recall that on $C_2(t, \theta)$, $\boldsymbol{\alpha} \cdot \eta = \boldsymbol{\alpha} \cdot (-1, 0) = (\alpha_1(t_1) - t_1) f(\mathbf{t})$, and therefore

$$
-\int_{T(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} = (\alpha_1(t_1) - t_1) \int_{C_2(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} + \int_{C_1(t_1,\theta)} \phi(\mathbf{t}) \cdot \eta(\mathbf{t})\ \mathrm{d}\mathbf{t}.
$$

Therefore,

$$\begin{aligned}\int_{C_1(t_1,\theta)} \phi(\mathbf{t})\cdot\eta(\mathbf{t})\ \mathrm{d}\mathbf{t} &= (t_1-\alpha_1(t_1))\int_{C_2(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} - \int_{T(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}\\ &= [\frac{\int_{T(t_1,1)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}}{\int_{C_2(t_1,1)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}}]\int_{C_2(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t} - \int_{T(t_1,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}, \qquad (5)\end{aligned}$$

directly from the definition of α_1.

We will now finish the proof that $\alpha_2(\mathbf{t}) \le \alpha_1(\mathbf{t})$. The idea is that $(\alpha_1,\alpha_2)\cdot(-\theta,1)$ is roughly (up to normalization) the derivative, with respect to t_1, of the following integral

$$\int_{C_1(t_1,\theta)} \phi(\mathbf{t})\cdot\eta(\mathbf{t})\ \mathrm{d}\mathbf{t}.$$

We use Equation (5) for the above integral, and differentiate it to get $(\alpha_1,\alpha_2)\cdot(-\theta,1)$. In particular, for all t_1 and θ we can write,

$$\begin{aligned}[\alpha_1(t_1,t_1\theta),\alpha_2(t_1,t_1\theta)]\cdot(-\theta,1)f(t_1,t_1\theta) &= -\frac{d}{dt_1}\Big(\int_{x=t_1}^{1}[\alpha_1(x,x\theta),\alpha_2(x,x\theta)](-\theta,1)f(x,x\theta)\ \mathrm{d}x\Big)\\ &= \frac{d}{dt_1}\Big(\int_{x=t_1}^{1}[(x,x\theta)-\alpha_1(x,x\theta),\alpha_2(x,x\theta)](-\theta,1)f(x,x\theta)\ \mathrm{d}x\Big)\\ &= \frac{d}{dt_1}\Big(\int_{C_1(t_1,\theta)}(\mathbf{t}-(\alpha_1,\alpha_2)(\mathbf{t}))\cdot\eta(\mathbf{t})f(\mathbf{t})\ \mathrm{d}\mathbf{t}\Big)\\ &= \frac{d}{dt_1}\Big(\int_{C_1(t_1,\theta)}\boldsymbol{\alpha}(\mathbf{t})\cdot\eta(\mathbf{t})\ \mathrm{d}\mathbf{t}\Big).\end{aligned}$$

Now using Equation 5,

$$\begin{aligned}[\alpha_1(t_1,t_1\theta),\alpha_2(t_1,t_1\theta)]\cdot(-\theta,1)f(t_1,t_1\theta) &= \frac{d}{dt_1}\Big(\int_{T(t,1)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}\times\frac{\int_{y=0}^{t_1\theta} f(t_1,y)\ \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1,y)\ \mathrm{d}y} - \int_{T(t,\theta)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}\Big)\\ &= \int_{T(t,1)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}\times\frac{d}{dt}\Big(\frac{\int_{y=0}^{t_1\theta} f(t_1,y)\ \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1,y)\ \mathrm{d}y}\Big)\\ &\le 0,\end{aligned}$$

where the inequality holds by lemma 6. Since f is always non-negative, this proves that $\alpha_2(t_1,t_2) \le \frac{t_2}{t_1}\alpha_1(t_1,t_2)$.

Notice that by together with Equality 3, the above equation shows that for MR-independent distributions,

$$\alpha_2(t_1,t_2) = \frac{t_2}{t_1}\alpha_1(t_1,t_2).$$

□

Proof of Lemma 8. We need to prove

$$\bar{f}(t_1,\theta_1)\times\bar{f}(t_1',\theta_1')\le\bar{f}(t_1,\theta_1')\times\bar{f}(t_1',\theta_1),\qquad\forall t_1\le t_1',\theta_1\ge\theta_1'.$$

Recall that $\bar{f}(t_1,\theta_1)=f(t_1,t_1\theta_1)$. Since the distribution is a product one, this implies that $\bar{f}(t_1,\theta_1)=f_1(t_1)f_2(t_1\theta_1)$. Notice that pair of values $t\theta'$ and $t'\theta$ have the same geometric mean as the pair $t\theta$, $t'\theta'$. Also given the assumptions, $t\theta'\le t\theta, t'\theta'\le t'\theta$. GG-convexity implies that

$$f_2(t\theta)\times f_2(t'\theta')\le f_2(t\theta')\times f_2(t'\theta).$$

Multiplying both sides by $f_1(t_1)\times f_1(t_1')$ we get

$$f_1(t_1)f_2(t\theta)\times f_1(t_1')f_2(t'\theta')\le f_1(t_1)f_2(t\theta')\times f_1(t_1')f_2(t'\theta),$$

which since the distribution is a product distribution implies that

$$\bar{f}(t_1,\theta_1)\times\bar{f}(t_1',\theta_1')\le\bar{f}(t_1,\theta_1')\times\bar{f}(t_1',\theta_1).$$

□

B.2 Proofs from Section 6.3

Proof of Lemma 17. Since $\alpha_2(\mathbf{t})\le\frac{t_2}{t_1}\alpha_1(\mathbf{t})$, the revenue of an allocation function can be upper bounded as follows,

$$\int\int\mathbf{x}(\mathbf{t})\cdot(\alpha_1(\mathbf{t}),\alpha_2(\mathbf{t}))f(\mathbf{t})\,\mathrm{d}(\mathbf{t})\le\int\int\mathbf{x}(\mathbf{t})\cdot(1,\frac{t_2}{t_1})\alpha_1(\mathbf{t})f(\mathbf{t})\,\mathrm{d}\mathbf{t}$$

After a changing variables according to $\theta=\frac{t_2}{t_2}$, the upper bound becomes

$$\int_{t_1}\int_{\theta=0}^{1}\bar{\mathbf{x}}(t,\theta)\cdot(1,\theta)\bar{\alpha}(t,\theta)\bar{f}(t,\theta)t_1\,\mathrm{d}\theta\,\mathrm{d}t_1,$$

where $\bar{\alpha}(t,\theta)=\alpha(t,t\theta)$. We perform another change of variables according to $q=\int_{T=T(t,1)}f(\mathbf{t})\,\mathrm{d}\mathbf{t}$, and rewrite the upper bound as

$$\int_q\int_\theta\mathbf{x}(q,\theta)(1,\theta)\alpha(q)\,\mathrm{d}\theta\,\mathrm{d}q.$$

Using integration by parts,

$$\int_q\int_\theta\mathbf{x}(q,\theta)(1,\theta)\alpha(q)\,\mathrm{d}\theta\,\mathrm{d}q=\int_q\int_\theta\frac{d}{dq}\mathbf{x}(q,\theta)(1,\theta)R(q)\,\mathrm{d}\theta\,\mathrm{d}q$$

where $R(q)$ is the revenue of posting a price that is accepted with probability q. Now let $\hat{R}$ be the convex hull of function R. In particular, $\hat{R}'(q)$ is non-decreasing and $\hat{R}(q)\ge R(q)$. We next show that $\frac{d}{dq}\mathbf{x}(q,\theta)\cdot(1,\theta)\ge0$ for any incentive compatible allocation rule.

Recall that the incentive compatibility condition states that, for all $\mathbf{t}$, $\mathbf{t}'$,

$$(\mathbf{x}(\mathbf{t}') - \mathbf{x}(\mathbf{t})) \cdot (\mathbf{t}' - \mathbf{t}) \quad \geq \quad 0.$$

Setting $\mathbf{t}' = \mathbf{t}(1+\epsilon)$, this becomes

$$(\mathbf{x}(\mathbf{t}(1+\epsilon)) - \mathbf{x}(\mathbf{t})) \cdot (\mathbf{t}\epsilon) \quad \geq \quad 0.$$

When ϵ goes to zero, this give us the following property

$$\frac{d}{dt_1}\mathbf{x}(\mathbf{t}) \cdot \mathbf{t} \geq 0.$$

Now notice that $q = \int_{T=T(t,1)} f(\mathbf{t})\ \mathrm{d}\mathbf{t}$ is a monotone function of t_1, which implies the desired property

$$\frac{d}{dq}\mathbf{x}(q,\theta) \cdot (1,\theta) \geq 0.$$

The facts that $\frac{d}{dq}\mathbf{x}(q,\theta) \cdot (1,\theta) \geq 0$ and $\hat{R}(q) \geq R(q)$ imply that we can upper bound revenue as follows

$$\int_q \int_\theta \frac{d}{dq}\mathbf{x}(q,\theta)(1,\theta)R(q)\ \mathrm{d}\theta\ \mathrm{d}q \leq \int_q \int_\theta \frac{d}{dq}\mathbf{x}(q,\theta)(1,\theta)\hat{R}(q)\ \mathrm{d}\theta\ \mathrm{d}q$$

We now study the allocation rule that optimizes virtual value function $(1,\theta)R'(q)$. Notice that when $\hat{R}(q) > R(q)$, we have that $\hat{R}'(q)$ is a constant and therefore $\frac{d}{dq}\mathbf{x} = 0$. This implies that $\hat{R}'(q)$ can be used as an upper bound on revenue, and for the expected virtual value of the optimal allocation exactly equals revenue. □