Reverse Mechanism Design

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

Jason Hartline
Northwestern University
EECS Department
hartline@northwestern.edu

May 23, 2014

Abstract

Myerson's 1981 characterization of revenue-optimal auctions for single-dimensional agents follows from an amortized analysis of the revenue from 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 and then identify sufficient conditions for its optimality. 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 distributions over types. 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). Similarly to an approach of Alaei et al. (2013), these results for single-agent pricing problems described above can be generalized 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 application. The consequence of our work is similar in that we characterize simple optimal mechanisms for well-behaved preferences; but distinct in that it does not characterize optimal mechanisms beyond the class of well-behave 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 global objective optimizing revenue in expectation over the distribution of agent values to the pointwise objective of optimizing virtual surplus on each profile of agent values. Furthermore, virtual surplus maximization leads to a 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 pref-

erences, 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 agents' incentives, it cannot be optimized pointwise. To illustrate this fact, consider a single agent wishing to buy a single item and whose value for the item is 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 an expected revenue of 0.21, and in the second mechanism at a price of 0.5 with probability 0.5 for an expected 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. The following two properties collectively imply that the mechanism that optimizes virtual values is indeed the optimal mechanism.[1] A *virtual value function* maps values pointwise to virtual values satisfying two properties:

- The pointwise optimization of virtual values gives an allocation rule that is *incentive compatible*. That is, there exist payments for this allocation rule that induce an agent to truthfully report his value.
- The virtual values are an *amortization* of the revenue. That is, the expected sum of the virtual values of the winners of any incentive compatible mechanism is equal to the expected revenue of that mechanism.

This 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 identification 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 (for a single agent) 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: serving the agent if her value is above 0.5 is incentive

[1] Subsequently, in Section 3, we will define a more permissive version of the amortization property which corresponds to the *ironed virtual values* of Myerson (1981). The simpler definition here will, nonetheless, be sufficient for our introductory discussion.

compatible and the appropriate payment is 0.5 if the agent is served and zero otherwise. This virtual value function satisfies the amortization condition: for posting any price p, 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 $p\,(1-p)$.[2]

The Challenge of Multi-dimensional Preferences. As described above, if a virtual value function that satisfies properties the incentive compatibility and amortization conditions can be identified then the optimal mechanism design problem is solved. For agents with single-dimensional preferences the function that satisfies the amortization property 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 (and can be found via integration by parts). It then remains to check the incentive compatibility property, i.e., that pointwise optimization of virtual value is incentive compatible; by standard characterizations of incentive compatible mechanisms, this is a simple task as well. Multi-dimensional preferences are challenging because the amortization property 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 amortization property 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 no item if 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 no item if all virtual values are negative. Synthesizing these constraints, the following conditions are sufficient for virtual surplus maximization to imply optimality of uniform pricing.

- The virtual value function is a *single-dimensional projection* if the virtual value for the favorite item corresponds to the single-dimensional virtual value for the distribution of the value for the favorite item (the distribution of the maximum value).[3]
- The virtual value function is *consistent with uniform pricing* if there is a price such that (a) when the value for the favorite item exceeds the price the virtual value for the favorite item is non-negative and at least the virtual value of any other items and (b) when the value for the favorite item is below the price both virtual values are non-positive.

Any virtual value function that satisfies the consistency-with-uniform-pricing conditions can be easily seen to satisfy the incentive compatibility requirement. Thus, the identification of a virtual value function becomes one of simultaneously resolving the three conditions of amortization,

[2] Further discussion omitted, this amortized equivalence continues to hold for randomized mechanisms.

[3] Rationale: 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 is the one given by the single-dimensional virtual values of Myerson (1981)

single-dimensional projection, and consistency with uniform pricing (which implies incentive compatibility). Conditions on the distributions over values for that guarantee the existence of such a virtual value function are sufficient for the optimality of uniform pricing.

For two-dimensional preferences the amortization and single-dimensional-projection restrictions pin 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 amortization condition are unique. Our task is then to give sufficient conditions under which this unique virtual value function is consistent with uniform pricing. 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 such virtual value function that satisfied all three constraints.

The restriction of selling a unit-demand agent his favorite item provides our main example of the framework of reverse mechanism design. We will also apply the framework to agents with additive preferences and give sufficient conditions for pricing the grand bundle to be optimal. The techniques we develop can be similarly applied to other environments and appropriate restrictions, and these extensions are an important topic for future work.

Sufficient Conditions for Optimality of Simple 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*. Such correlation is natural when values correlate, e.g., with initial wealth. Below we state the sufficient conditions for optimality of uniform pricing for unit-demand agents and grand-bundle pricing for additive agents for the special case of two items; the general multi-item statements are deferred to later in the paper.

A single-agent problem with unit-demand preferences is defined with a density function f on a bounded set of types normalized to $T = [0, 1]^2$, where (t_1, t_2) is the pair of values for the two items. We consider only symmetric distributions and therefore only define f on the subset of T satisfying $t_1 \geq t_2$, that is, where the first item is the favorite item. A uniform posted price mechanism posts a uniform price for each item, and lets the agent choose one of the two items or nothing based on his type. It is easiest to state the result in the following *max-ratio* representation of the distribution. The max-ratio representation of a density function f is the function $\bar{f}$ that is defined by

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

The following theorem states that uniform pricing is optimal if $\bar{f}$ satisfies a supermodularity condition. Intuitively, this supermodularity implies that a higher value of t_1 signals a higher value of t_2/t_1 (see Figure 1).

Theorem 1. *Posting a uniform price for each item is optimal if the max-ratio representation of the density function is log-supermodular, that is,*

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

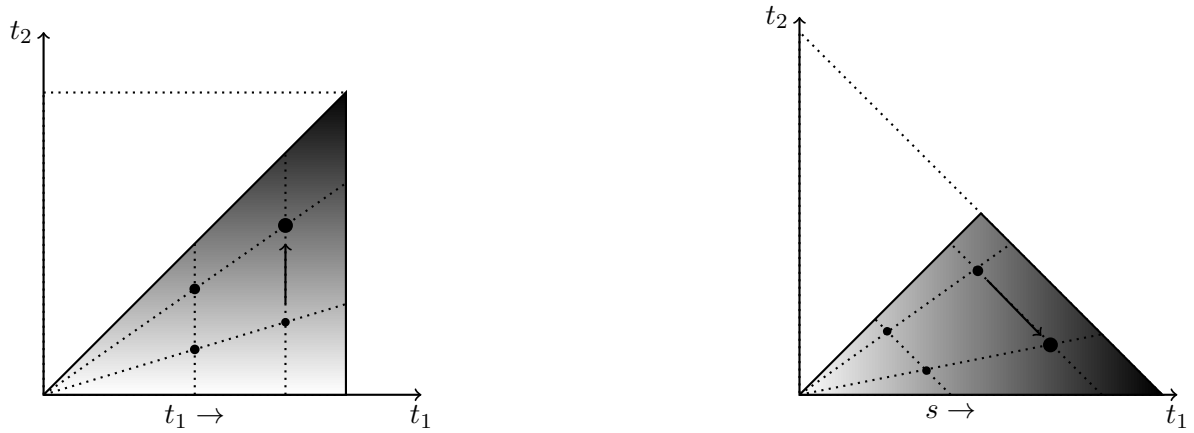


Figure 1: a) A distribution satisfying the supermodularity condition of Theorem 1. Dark color indicates relatively high mass. Roughly speaking, it states that the probability mass shifts upwards as the value for the favorite item, t_1, increases. b) A distribution satisfying the submodularity condition of Theorem 2. It implies that probability mass shifts downwards as the sum of the values increases.

As a very simple example, for the two-dimensional case, consider selling two forms of energy, e.g., electricity and gas, to a plant. The plant uses electricity, and its value for a unit of electricity, v, is uniform from interval $[0, 1]$. The plant is capable of transforming a unit of gas to electricity using a generator with efficiency $\theta \in [0, 1]$, drawn independently of v. The value for a unit of gas is therefore θv where both v and θ are private to the plant. Our theorem states that the optimal auction is to simply post a price for a unit of electricity. The two dimensional the virtual value function that satisfies the amortization and single-dimensional-projection constraints is pinned down 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 are consistent with uniform pricing. Therefore, the existence of this virtual value function proves that uniform pricing is revenue optimal among all mechanisms.

For other i.i.d. distributions, the virtual values that amortization and single-dimensional-projection pin down do not generally satisfy consistency with uniform pricing (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 consistency with uniform pricing is satisfied by for multi-dimensional values that are i.i.d. from any distribution that satisfies a geometric convexity property.

As the second main example of the application of our framework, we provide sufficient conditions for optimality of grand bundle pricing for agents with additive preferences. The *sum-ratio* representation of a density function f is the function $\bar{f}$ that is defined by

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

Theorem 2. *Posting a price for the bundle of items is optimal if the sum-ratio representation of the density function is log-submodular, that is,*

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

Related Work. The starting point of work in multi-dimensional optimal mechanism design is 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., McAfee and McMillan, 1988; Armstrong, 1996). The surplus 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., McAfee and McMillan, 1988; Manelli and Vincent, 2006; Hart and Nisan, 2013; Daskalakis et al., 2013; Wang and Tang, 2014; Giannakopoulos and Koutsoupias, 2014) or in terms of the gradient of the utility (e.g., Armstrong, 1996; Alaei et al., 2013; and this paper). This manipulation follows from an integration by parts. The first category of papers (rewriting objective in terms of utility) performs the integration by parts independently in each dimension, and the second category (rewriting objective in terms of gradient of utility, except for ours) does the integration along rays from the origin (see below). In our approach, in contrast, the integration by parts is performed in general and is dependent on the distribution and the form of the mechanism we wish to show is optimal.

Closest to our work are Armstrong (1996) and Alaei et al. (2013) which use integration by parts along paths that connect types with straight lines to the zero type (which has value zero for any outcome) to define virtual values. For the first work, Armstrong (1996) finds properties on valuation functions (beyond linear ones considered in our paper) and distributions, that when jointly satisfied, imply that the pointwise optimization of virtual values results in an incentive compatible mechanism. Armstrong gives some examples of mechanisms that result from this approach but does not generally interpret the form of the resulting mechanisms. Armstrong suggests generalizing his approach from rays from the origin to other kinds of paths; our approach, in contrast, proves the existence of appropriate paths over which to integrate without requiring the form of the path to be specified in advance. When Armstrong's condition on the distribution is satisfied (which we refer to as independence in max-ratio coordinates), our solution is also equivalent to an integration along rays. For the second work, 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 ours) when the single-agent mechanism design problems satisfy a *revenue linearity* property. They prove that a unit-demand agent with values for items that are independently, identically, and uniformly distributed on the $[0, 1]$ interval is revenue linear; our results generalize this one. Moreover, 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 in the *service constrained environments* of Alaei et al. (2013).

There has been work looking at properties of single-agent mechanism design problems that are sufficient for optimal mechanisms to make only limited use of randomization. For context, the optimal single-item mechanism is always deterministic (e.g., Myerson, 1981; Riley and Zeckhauser, 1983), while the optimal multi-item mechanism is sometimes randomized (e.g., Thanassoulis, 2004; Pycia, 2006). For agents with additive preferences across multiple items, McAfee and McMillan (1988), Manelli and Vincent (2006), and Giannakopoulos and Koutsoupias (2014) find sufficient conditions under which deterministic mechanisms, i.e., bundle pricings, are optimal. Pavlov (2011) considers more general preferences and a more general condition; for unit-demand preferences, this condition implies that in the optimal mechanism an agent deterministically receives an item or not, though the item received may be randomized.

A number of papers consider the question of finding closed forms for the optimal mechanism for

an agent with additive preferences and independent values across the items. One such closed form is grand-bundle pricing. Our work for additive preferences contrasts in that we restrict to grand-bundle pricing and a particular family of correlated distributions. For the two item case, Hart and Nisan (2013) give sufficient conditions for the optimality of grand-bundle pricing; these conditions are further generalized by Wang and Tang (2014). Their results are not directly comparable to ours as our results apply to correlated distributions. Daskalakis et al. (2013) and Giannakopoulos and Koutsoupias (2014) give frameworks, similar to ours, for proving optimality of multi-dimensional mechanisms. Daskalakis et al. (2013) use their framework to find the close form for the optimal mechanism for several special cases; Giannakopoulos and Koutsoupias (2014) give a closed form for the optimal mechanism when values are i.i.d. from the uniform distribution (with up to six items).

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 $\boldsymbol{t} = (t_1, \ldots, t_m) \in T$ is an m-dimensional vector of values for m items.[4] The type of the agent is drawn from a known distribution with density f. For the special case that the type space is single-dimensional (i.e., $m = 1$), the cumulative distribution function of the type is denoted by F. We do not require that the values for items be drawn independently. We consider unit-demand and additive agents. For unit-demand agents the allocation $\boldsymbol{x} \in [0,1]^m$ must satisfy $\sum_i x_i \leq 1$; for additive agents $\boldsymbol{x}$ must satisfy $x_i \leq 1$ for all i. The utility of the agent with type $\boldsymbol{t}$ for allocation $\boldsymbol{x} \in [0,1]^m$ and payment $p \in \mathbb{R}$ is $\boldsymbol{t} \cdot \boldsymbol{x} - p$.

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

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

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

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

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

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

[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.

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

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

Let ∂T be the boundary of set T and $\boldsymbol{\eta}(\boldsymbol{t})$ be the outward-pointing unit normal vector of T at point $\boldsymbol{t}$ on ∂T. The multi-variable *integration by parts* for functions $h : \mathbb{R}^k \to \mathbb{R}$ and $\boldsymbol{\alpha} : \mathbb{R}^k \to \mathbb{R}^k$ is as follows

$$\int_{\boldsymbol{t}\in T} (\nabla h \cdot \boldsymbol{\alpha})(\boldsymbol{t})\ \mathrm{d}\boldsymbol{t} = \int_{\boldsymbol{t}\in \partial T} h(\boldsymbol{t})(\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\boldsymbol{t})\ \mathrm{d}\boldsymbol{t} - \int_{\boldsymbol{t}\in T} h(\boldsymbol{t})(\nabla\cdot\boldsymbol{\alpha}(\boldsymbol{t}))\ \mathrm{d}\boldsymbol{t}, \tag{1}$$

Setting h to be the constant function equal to 1 everywhere gives us the *divergence theorem*

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

When the dimension $k = 1$, 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 is 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 $(\boldsymbol{x}, p)$ defines a utility function $u(\boldsymbol{t}) = \boldsymbol{t}\cdot\boldsymbol{x}(\boldsymbol{t}) - p(\boldsymbol{t})$. The following lemma connects the utility function of an IC mechanism with its allocation function.

Lemma 3 (Rochet, 1985)**.** *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 $\boldsymbol{x}(\boldsymbol{t}) = \nabla u(\boldsymbol{t})$.*

Notice that the payment function can be defined using the utility function and the allocation function as $p(\boldsymbol{t}) = \boldsymbol{t}\cdot\boldsymbol{x}(\boldsymbol{t}) - u(\boldsymbol{t})$. Applying the above lemma, we can write payment to be $p(\boldsymbol{t}) = \boldsymbol{t}\cdot\nabla u(\boldsymbol{t}) - u(\boldsymbol{t})$. The revenue maximization problem can then be written as the following mathematical program, which is the starting point for the analysis of this paper.

$$\begin{aligned} \max_{\boldsymbol{x},u} \quad & \int_{\boldsymbol{t}} [\boldsymbol{t}\cdot\nabla u(\boldsymbol{t}) - u(\boldsymbol{t})] f(\boldsymbol{t})\ \mathrm{d}\boldsymbol{t} \\ & \boldsymbol{x} = \nabla u; u \text{ is convex}, \\ & \forall \boldsymbol{t}, \boldsymbol{x}(\boldsymbol{t}) \text{ is feasible allocation.} \end{aligned} \tag{2}$$

Notice that when the dimension of the type space is $m = 1$, the above program is equivalent to the following familiar form from Myerson (1981),

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

In Section 6, we extend the above formulation and our results to multi-agent settings.

3 Amortization of Revenue

This section formalizes our codification of multi-dimensional virtual values for incentive compatible mechanism design and describes the working pieces of our framework. The main construct is the definition of multi-dimensional virtual value functions and the accompanying proposition, below.

Definition 1. A *vector field* $\bar{\boldsymbol{\phi}} : [0,1]^m \to \mathbb{R}^m$ that maps an an m-dimensional type to an m-dimensional vector is

- *incentive compatible* if the *virtual surplus maximizer* given by selecting the outcome $\boldsymbol{x}$ for type $\boldsymbol{t}$ that optimizes virtual surplus $\boldsymbol{x} \cdot \bar{\boldsymbol{\phi}}(\boldsymbol{t})$ is incentive compatible;
- a *weak amortization of revenue* if, in expectation over types drawn from the distribution, the virtual surplus of any incentive compatible mechanism upper bounds is revenue, i.e., $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t})] \geq \mathbf{E}[p(\boldsymbol{t})]$, and with equality for the virtual surplus maximizer;
- a *strong amortization of revenue* if the inequality (of weak amortization) holds with equality for all incentive compatible mechanisms; and
- a *virtual value function* if it is incentive compatible and a weak or strong amortization.

Proposition 4. *In any environment for which a virtual value function exists, the virtual surplus maximizer is incentive compatible and revenue optimal.*

Proof. The expected revenue of the virtual surplus maximizer is equal to its expected virtual surplus (by weak amortization). This expected virtual surplus is at least the virtual surplus of any alternate mechanism (by definition of virtual surplus maximization). The expected virtual surplus is an upper bound on the expected revenue of the alternative mechanism (by weak amortization). Thus, the expected revenue of the virtual surplus maximizer is at least that of the alternative mechanism. Incentive compatibility follows directly from the definition of a virtual value function. □

For a single-dimensional agent (i.e., $m = 1$), Myerson (1981) showed that the function $v - \frac{1-F(v)}{f(v)}$ is a strong amortization, when it is monotone it is incentive compatible, when it is non-monotone an ironing procedure can be applied to obtain from it a weak amortization function that is monotone and thus incentive compatible. Our approach will analogously enable the derivation of multi-dimensional virtual value functions (i.e., satisfying incentive compatibility and weak amortization) via the construction of a strong amortization function that is not necessarily incentive compatible.

We now give sufficient conditions for a vector field to be a strong amortization of revenue. At a high level, we derive first a strong amortization of utility and then, using the fact that revenue is value minus utility, derive a strong amortization of revenue. These strong amortizations will be building blocks for the derivation of virtual values for unit-demand and additive agents in the subsequent sections. The following lemma follows from integration by parts as per equation (1), the definition of strong amortization of utility (Definition 1, generalized to utility), and the fact that the gradient of utility is the allocation rule of the mechanism (Lemma 3).

Lemma 5. *For type space T and distribution f, vector field $\boldsymbol{\alpha}/f$ is a strong amortization of utility, i.e., $\mathbf{E}[u(\boldsymbol{t})] = \mathbf{E}[\boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t})]$ for all incentive compatible allocation rules $\boldsymbol{x}$, if it satisfies*

- divergence density equality*, i.e., that $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}) = -f(\boldsymbol{t})$ at any point $\boldsymbol{t} \in T$, and*

- boundary orthogonality, *i.e., that* $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) = 0$ *for all* $\boldsymbol{t} \neq \mathbf{0}$ *on the boundary of type space* ∂T *with normal vector* $\boldsymbol{\eta}(\boldsymbol{t})$.

Proof. Write the expectation of $\mathbf{E}[\boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t})]$ as the integral $\int_{\boldsymbol{t} \in T} \boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}$. From Lemma 3, substitute ∇u for allocation $\boldsymbol{x}$ and apply integration by parts.

$$\int_{\boldsymbol{t} \in T} \boldsymbol{\alpha}(\boldsymbol{t}) \cdot \nabla u(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} = \int_{\boldsymbol{t} \in \partial T} u(\boldsymbol{t}) (\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} - \int_{\boldsymbol{t} \in T} u(\boldsymbol{t}) \, (\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t})) \, \mathrm{d}\boldsymbol{t} \tag{3}$$

$$= u(\boldsymbol{t}) + \int_{\boldsymbol{t} \in T} u(\boldsymbol{t}) \, f(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} \tag{4}$$

The second equality is derived from the first equality by employing the assumptions of the lemma on vector field $\boldsymbol{\alpha}$ as follows.

- By divergence density equality, the second term simplifies by substituting $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}) = -f(\boldsymbol{t})$.
- Recall that the divergence theorem is equivalent setting $u(\boldsymbol{t}) = 1$ in the formula (3), this gives

$$\int_{\boldsymbol{t} \in \partial T} (\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} = \int_{\boldsymbol{t} \in T} f(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} = 1, \tag{5}$$

  the total probability of any type. Boundary orthogonality implies that the integrand in the boundary integral of equation (5) is identically zero everywhere except $\boldsymbol{t} = 0$. To integrate to one on the boundary, the function must be the Dirac delta function at $\boldsymbol{t} = 0$; thus, the integral of the first term in equation (3) is $u(\mathbf{0})$.

Without loss of generality for revenue optimal mechanisms $u(\mathbf{0}) = 0$. We can interpret the left- and right-hand sides of equation (4) as expectations, which gives $\mathbf{E}[\boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t})] = \mathbf{E}[u(\boldsymbol{t})]$, the definition of strong amortization of utility for $\boldsymbol{\alpha}/f$. □

For a single-dimensional agent with value v in type space $T = [0, 1]$, the only function that satisfies the conditions of Lemma 5 and gives a strong amortization of utility is $\alpha(v)/f(v) = \frac{1-F(v)}{f(v)}$. For this formula, notice that the divergence of $\alpha(v) = 1 - F(v)$ is simply its derivative $-f(v)$. The boundary $\partial T \setminus \{0\}$ is the point $v = 1$, the upper bound of the distribution, and thus trivially satisfies orthogonality as $\alpha(1) = 0$. In classical auction theory the amortization of utility $\frac{1-F(v)}{f(v)}$ is often referred to as the agent's *information rent*.

The following lemma is immediate from the fact that revenue is the agent's surplus minus the agent's utility. For a single-dimensional agent it implies that $\boldsymbol{\phi}(v) = v - \frac{(1-F(v))}{f(v)}$ is the strong amortization of revenue.

Lemma 6. *For type space* T *and distribution* f*, vector field* $\boldsymbol{\phi}$ *is a strong amortization of revenue, i.e.,* $\mathbf{E}[p(\boldsymbol{t})] = \mathbf{E}[\boldsymbol{\phi}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t})]$ *for all incentive compatible allocation rules* $\boldsymbol{x}$*, if and only if* $\boldsymbol{\phi}(\boldsymbol{t}) = \boldsymbol{t} - \boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t})$ *for all* $\boldsymbol{t}$ *but a measure zero subset of* T *and* $\boldsymbol{\alpha}/f$ *is a strong amortization of utility.*

Unlike the case of a single-dimensional agent, for multi-dimensional agents there are many strong amortizations of utility and, consequentally, many strong amortizations of revenue. As an example, suppose we wish to show the optimality of a restricted form of mechanism via a strongly amortized virtual value function for an $m = 2$ dimensional agent. This virtual value fuction has two degrees of freedom. We can pin down one degree of freedom by equating virtual surplus to

expected revenue for mechanisms with this restricted form. The divergence density equality for strong amortizations that Lemma 6 inherits from Lemma 5 gives a differential equation that then pins down the other degree of freedom. It remains to find sufficient conditions on the distribution under which virtual surplus maximization identically gives mechanisms of the restricted form.

The approach above can be generalized to show optimality of mechanisms via a weakly amortized vitual value function. For example, such a generalization can be used to give proofs of optimality under more permissive distributional assumptions. To subsitute weak amortization for stong amortization we need a way to relate the differential equations (from divergence density equality) that govern strong amortizations to any given weak amortization. Such a relationship follows directly from the definitions of both weak and strong amortization in terms of the expected revenue of any incentive compatible mechanism (Definition 1) and is summarized below as Lemma 7.

Lemma 7. *For type space T and distribution f, vector field $\bar{\boldsymbol{\phi}}$ is a weak amortization of revenue if and only if there exists a strong amortization of revenue $\boldsymbol{\phi}$ such that $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t})\,\boldsymbol{x}(\boldsymbol{t})] \geq \mathbf{E}[\boldsymbol{\phi}(\boldsymbol{t})\,\boldsymbol{x}(\boldsymbol{t})]$, for all incentive compatible mechanisms $\boldsymbol{x}$, with equality for $\boldsymbol{x}$ that pointwise maximizes $\bar{\boldsymbol{\phi}}$.*

4 Optimality of Uniform Pricing for Unit Demand Preferences

In this section we study the existence virtual values to prove optimality of uniform pricing for a single unit-demand agent. To simplify exposition we focus on the case of two items and on the type space $T = \{\boldsymbol{t} = (t_1, t_2) : 0 \leq t_2 \leq t_1 \leq 1\}$. In this type space item one is always the agent's favorite item. Our conclusions easily to the $[0,1]^2$ type space with symmetric disitributions; other extensions are given at the end of the section. The general case of $m \geq 2$ items is considered in Appendix A.

The following outline describes the steps to be taken for showing optimality of uniform pricing via a strongly amortized virtual value function.

1. Single-dimensional projection (for uniform pricing): By the definition of virtual value functions, it must be that the expected revenue from posting a uniform price is equal to its expected virtual surplus. From this constraint, we will derive a formula for $\phi_1(\boldsymbol{t})$. From Lemma 6 and this formula for $\phi_1(\boldsymbol{t})$, we obtain a formula for $\alpha_1(\boldsymbol{t})$ as $(t_1 - \phi_1(\boldsymbol{t}))f(\boldsymbol{t})$.

2. Divergence density equality: The divergence density equality of Lemma 5 relates $\alpha_1(\boldsymbol{t})$ to $\alpha_2(\boldsymbol{t})$ and from it we obtain an equation for $\alpha_2(\boldsymbol{t})$. From Lemma 6 and this formula for $\alpha_2(\boldsymbol{t})$, we obtain a formula for $\phi_2(\boldsymbol{t})$ as $t_2 + \phi_2(\boldsymbol{t})/f(\boldsymbol{t})$.

3. Boundary orthogonality: For the identified $\boldsymbol{\alpha}$ we verify that boundary orthogonality (required by Lemma 5) is satisfied. For the triangular shaped type space T_1 this requires checking orthoganlity of $\boldsymbol{\alpha}$ to the three edges of the triangle.

4. Consistency with uniform pricing: We identify conditions on the distribution f such that the virtual surplus maximization (with respect to $\boldsymbol{\phi}$) corresponds to uniform pricing. This consistency implies that the strong amortization given by $\boldsymbol{\phi}$ is incentive compatible.

5. Uniform pricing is optimal: From the derivation above and Proposition 4, under the distributional conditions identified, uniform pricing is optimal.

6. Ironing for weak amortization. We will show in Section 6.3 how the assumption that the single-dimensional projection has a strong amortization which is monotone can be relaxed with a technique similar to Myerson's.

For $0 \le v \le 1$, define $F_{\max}(v) = \int_{\boldsymbol{t} \in T, t_1, t_2 \le v} f \, \mathrm{d}\boldsymbol{t}$ to be the distribution of the agent's favorite item, and $f_{\max}$ to be the corresponding density function.

Lemma 8 (single-dimensional projection)**.** *The expected revenue of a uniform pricing is equal to its expected virtual surplus with respect to vector field $\boldsymbol{\phi}$ with $\phi_1(\boldsymbol{t}) = \phi_{\max}(t_1)$ where $\phi_{\max}(v) = v - \frac{1-F_{\max}(v)}{f_{\max}(v)}$ is the single-dimensional strong amortization for agent's favorite item.*

Proof. When posting price p for either item, the agent always only buys his favorite item. Thus, uniform pricing only ever sells the favorite item. By the single-dimensional theory of optimal auctions, the expected revenue of any mechanism for selling only a single item is equal to its expected virtual surplus according to a single-dimensional strong amortization. As the uniform pricing only sells the favorite item its multi-dimensional virtual surplus is equal to its virtual surplus of the favorite item. As the statement of the lemma defines the virtual value of the favorite item to be equal to the single-dimensional strong amortization, these multi-dimensional and single-dimensional virtual surpluses of uniform pricing are equal. Thus, the expected multi-dimensional virtual surplus is also equal to the expected revenue. The specific form of the single-dimensional strong amortization given in the lemma is standard from Myerson (1981) (see Appendix B for a more formal proof of this lemma). □

For $\phi_1(\boldsymbol{t})$, defined above, to be the first coordinate of a strong amortization of revenue, there must be a corresponding strong amortization of utility $\boldsymbol{\alpha}/f$. Lemma 6 gives the relationship between these strong amortizations as $\boldsymbol{\alpha}(\boldsymbol{t}) = (\boldsymbol{t} - \boldsymbol{\phi}(\boldsymbol{t}))\, f(\boldsymbol{t})$. We have identified $\phi_1(\boldsymbol{t})$ above, from it we have $\alpha_1(\boldsymbol{t})$. The lemma below, uses the divergence density equality of $\boldsymbol{\alpha}$ to specify $\alpha_2(\boldsymbol{t})$, from which $\phi_2(\boldsymbol{t})$ can be derived.

Lemma 9 (divergence density equality)**.** *A strong amortization of utility $\boldsymbol{\alpha}$ satisfies*

$$\alpha_2(t_1, t_2) = -\int_{y=0}^{t_2} f(t_1, y) + \partial_1 \alpha_1(t_1, y)) \, \mathrm{d}y.$$

Proof. From the right-hand side of the statement of the lemma we have,

$$\begin{aligned} \partial_2 \alpha_2(t_1, t_2) &= \frac{\mathrm{d}}{\mathrm{d}t_2}[\alpha_2(t_1, t_2))] \\ &= -f(t_1, t_2) - \partial_1 \alpha_1(t_1, t_2), \end{aligned}$$

and as a result,

$$\begin{aligned} \nabla \cdot \boldsymbol{\alpha} &= \partial_1 \alpha_1(t_1, t_2) + \partial_2 \alpha_2(t_1, t_2) \\ &= -f(t), \end{aligned}$$

as required by divergence density equality (of Lemma 5). □

Lemma 10 (boundary orthogonality)**.** *For $\boldsymbol{\alpha}$ defined from ϕ_1 from the single-dimensional projection to the favorite item, boundary orthogonality is satisfied.*

Proof. We must show that the vector field $\boldsymbol{\alpha}$ is orthogonal to the boundary of T. As T is given by a triangle, the argument decomposes into analysis of the three edges of the triangle.

- Consider the $t_1 = 1$ edge of T. By the single-dimensional projection, $\phi_1(1, t_2) = \phi_{\max}(1) = 1 - \frac{1-F_{\max}(1)}{f_{\max}(1)} = 1$ (as $F_{\max}(1) = 1$). So $\alpha_1(1, t_2) = (1 - \phi_{\max}(1))f(1, t_2) = 0$. Thus, $\boldsymbol{\alpha}$ is orthogonal to the normal vector $\boldsymbol{\eta}(1, t_2) = (1, 0)$.

- Consider the $t_2 = 0$ edge of T. By Lemma 9, $\alpha_2(\boldsymbol{t})$ is given by an integral from zero to t_2. With $t_2 = 0$, $\alpha_2(t_1, 0) = 0$. Thus, $\boldsymbol{\alpha}$ is orthogonal to the normal vector $\boldsymbol{\eta}(t_1, 0) = (0, -1)$.

- Consider the $t_1 = t_2$ edge of T. Orthoganality of $\boldsymbol{\alpha}$ to normal $\boldsymbol{\eta}(t_1, t_1) = (-1, 1)/\sqrt{2}$ requires $\alpha_1(t_1, t_1) = \alpha_2(t_1, t_1)$ for any t_1. (Of course, symmetry requires this too.) By definition of α_2 (from Lemma 9),

$$\begin{aligned}
\int_{z=t_1}^{1} \alpha_2(z, z)dz &= -\int_{z=t_1}^{1} \int_{y=0}^{z} f(z, y) + \partial_1\alpha_1(z, y)) \, \mathrm{d}y\mathrm{d}z \\
&= -(1 - F_{max}(t_1)) - \int_{z=t_1}^{1} \int_{y=0}^{z} \partial_1\alpha_1(z, y) \, \mathrm{d}y\mathrm{d}z \\
&= -(1 - F_{max}(t_1)) - \int_{y=0}^{t_1} \int_{z=t_1}^{1} \partial_1\alpha_1(z, y) \, \mathrm{d}z\mathrm{d}y - \int_{y=t_1}^{1} \int_{z=y}^{1} \partial_1\alpha_1(z, y) \, \mathrm{d}z\mathrm{d}y.
\end{aligned}$$

Since $\alpha_1(1, t_2) = 0$,

$$= -(1 - F_{max}(t_1)) + \int_{y=0}^{t_1} \alpha_1(t_1, y) \, \mathrm{d}y + \int_{y=t_1}^{1} \alpha_1(y, y) \, \mathrm{d}y.$$

By definition, $\alpha_1(t_1, y) = \frac{1-F_{\max}(t_1)}{f_{\max}(t_1)} f(t_1, y)$, and thus we can write

$$\begin{aligned}
&= -(1 - F_{max}(t_1)) + \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} \int_{y=0}^{t_1} f(t_1, y) \, \mathrm{d}y + \int_{y=t_1}^{1} \alpha_1(y, y) \, \mathrm{d}y \\
&= -(1 - F_{max}(t_1)) + \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} f_{\max}(t_1) + \int_{y=t_1}^{1} \alpha_1(y, y) \, \mathrm{d}y \\
&= \int_{y=t_1}^{1} \alpha_1(y, y).
\end{aligned}$$

Since the two integrals are equal for all t_1, we must have $\alpha_1(t_1, t_1) = \alpha_2(t_1, t_1)$ for all t_1.

□

Lemma 11 (consistency with virtual surplus maximization; incentive compatibility)**.** *Virtual surplus maximization according to vector field $\boldsymbol{\phi}$ gives a uniform pricing p^* (and is incentive compatible) if and only if*

(a) $\phi_1(\boldsymbol{t}) \geq 0$ when $t_1 \geq p^*$ and $\phi_1(\boldsymbol{t}) \leq 0$ when $t_1 \leq p^*$, and

(b) $\phi_2(\boldsymbol{t}) \leq \phi_1(\boldsymbol{t})$ when either are positive.

Proof. We need to show that for the uniform price p, the allocation function $\boldsymbol{x}^p$ (defined in the proof of Lemma 8) optimizes $\boldsymbol{\phi}$ pointwise. Pointwise optimization of $\boldsymbol{x} \cdot \boldsymbol{\phi}$ will result in $\boldsymbol{x} = (1, 0)$ whenever $\phi_1 \geq 0$ and $\phi_1 \geq \phi_2$, and will result in $\boldsymbol{x} = (0, 0)$ whenever $\phi_1, \phi_2 \leq 0$. It is sufficient to have $\phi_1(\boldsymbol{t}) \geq \phi_2(\boldsymbol{t})$ whenever $t_1 \geq p^*$, and $\phi_2(\boldsymbol{t}) \leq 0$ whenever $t_1 \leq p^*$. □

We now give sufficient conditions on the distribution such that the strong amortization identified by the single-dimensional projection of Lemma 8, Lemma 9, and the mapping from α_2 to ϕ_2 given by Lemma 6) satisfies the conditions of Lemma 11. To specify these conditions, it is useful to view the distribution in the *max-ratio* (MR) coordinates. In particular the max-ratio representation f^{MR} of a density function f is defined to be

$$f^{MR}(t_1, \theta) = f(t_1, \theta t_1), \qquad \forall t_1, \theta, 0 \leq t_1, \theta \leq 1.$$

Equivalantly,

$$f(t_1, t_2) = f^{MR}(t_1, t_2/t_1) \qquad \forall t_1, t_2, 0 \leq t_1, t_2 \leq 1.$$

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

$$f^{MR}(t_1, \theta) \times f^{MR}(t_1', \theta') \leq f^{MR}(t_1, \theta') \times f^{MR}(t_1', \theta), \qquad \forall t_1 \leq t_1', \theta \geq \theta'.$$

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

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

We now show that MR-log-supermodularity is sufficient the strong amortization of revenue constructed from the single-dimensional projection satisfies Lemma 11 (maximizing virtual surplus is consistent with uniform pricing). This lemma combined with Proposition 4 gives the subsequent theorem.

Lemma 12. *If a distribution is MR-log-supermodular, then $\phi_2(t_1, t_2) \leq \frac{t_2}{t_1}\phi_1(t_1, t_2)$, with equality if the distribution is MR-independent.*

Proof. By definition of α_2 (from Lemma 9), and a series of calculations similar to the proof of Lemma 10 we have for all $t_1, \theta \leq 1$,

$$\begin{aligned} \int_{z=t_1}^{1} \alpha_2(z, \theta z) dz &= -\int_{z=t_1}^{1} \int_{y=0}^{z\theta} f(z, y) + \partial_1 \alpha_1(z, y))\ \mathrm{d}y\mathrm{d}z \\ &= -\int_{z=t_1}^{1} \int_{y=0}^{z\theta} f(z, y)\ \mathrm{d}y\mathrm{d}z - \int_{z=t_1}^{1} \int_{y=0}^{z\theta} \partial_1 \alpha_1(z, y)\ \mathrm{d}y\mathrm{d}z \\ &= -\int_{z=t_1}^{1} \int_{y=0}^{z\theta} f(z, y)\ \mathrm{d}y\mathrm{d}z - \int_{y=0}^{t_1\theta} \int_{z=t_1}^{1} \partial_1 \alpha_1(z, y)\ \mathrm{d}z\mathrm{d}y - \int_{y=t_1\theta}^{\theta} \int_{z=y/\theta}^{1} \partial_1 \alpha_1(z, y)\ \mathrm{d}z\mathrm{d}y. \end{aligned}$$

Since $\alpha_1(1, t_2) = 0$,

$$= -\int_{z=t_1}^{1} \int_{y=0}^{z\theta} f(z, y)\ \mathrm{d}y\mathrm{d}z + \int_{y=0}^{t_1\theta} \alpha_1(t_1, y)\ \mathrm{d}y + \int_{y=t_1\theta}^{\theta} \alpha_1(y/\theta, y)\ \mathrm{d}y.$$

By definition, $\alpha_1(t_1, y) = \frac{1-F_{\max}(t_1)}{f_{\max}(t_1)} f(t_1, y)$, and thus we can write

$$= -\int_{z=t_1}^{1} \int_{y=0}^{z\theta} f(z, y)\ \mathrm{d}y\mathrm{d}z + \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} \int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y + \int_{z=t_1}^{1} \alpha_1(z, z\theta)\theta\ \mathrm{d}z.$$

Using the above equation, we can write

$$\begin{aligned}
\alpha_1(t_1, t_1\theta)\theta - \alpha_2(t_1, t_1\theta) &= -\frac{\mathrm{d}}{\mathrm{d}t_1}[\int_{z=t_1}^{1} \alpha_1(z, z\theta)\theta - \alpha_2(z, z\theta)\ \mathrm{d}z] \\
&= -\frac{\mathrm{d}}{\mathrm{d}t_1}\left[\int_{z=t_1}^{1} \int_{y=0}^{z\theta} f(z, y)\ \mathrm{d}y\mathrm{d}z - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} \int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y\right] \\
&= \int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y + (1 - F_{\max}(t_1))\frac{\mathrm{d}}{\mathrm{d}t_1}\Big[\frac{\int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y}{f_{\max}(t_1)}\Big] - f_{\max}(t_1)\frac{\int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y}{f_{\max}(t_1)} \\
&= (1 - F_{\max}(t_1))\frac{\mathrm{d}}{\mathrm{d}t_1}\Big[\frac{\int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y}{f_{\max}(t_1)}\Big].
\end{aligned}$$

Lemma 15 shows that

$$\frac{\mathrm{d}}{\mathrm{d}t_1}\Big[\frac{\int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y}{f_{\max}(t_1)}\Big] \leq 0, \tag{6}$$

(with equality for MR-independent distributions) for MR-log-supermodular distributions (notice that $1 - F_{\max}(t_1)$ is non-negative), and therefore $\alpha_1(t_1, t_1\theta)\theta \leq \alpha_2(t_1, t_1\theta)$. Using the definitions, and by setting $\theta = t_2/t_1$,

$$\begin{aligned}
\frac{t_2}{t_1}\phi_1(\boldsymbol{t}) &= \frac{t_2}{t_1}(t_1 - \frac{\alpha_1(\boldsymbol{t})}{f(\boldsymbol{t})}) \\
&= t_2 - \frac{t_2}{t_1}\frac{\alpha_1(\boldsymbol{t})}{f(\boldsymbol{t})} \\
&\geq t_2 - \frac{\alpha_2(\boldsymbol{t})}{f(\boldsymbol{t})} \\
&= \phi_2(\boldsymbol{t}).
\end{aligned}$$

□

Theorem 13. *Uniform pricing is the optimal auction if the distribution is symmetric and max-ratio log-supermodular, and the strong amortization of revenue $\phi_{\max}(\cdot)$ for the single-dimensional projection is monotone.*

The conditions of Theorem 13 can be further relaxed by constructing a weak amortization $\bar{\boldsymbol{\phi}}$ from the strong amortization $\boldsymbol{\phi}$, above. The following is a special case of the more general Lemma 25 given in Section 6.3.

Theorem 14. *Uniform pricing is the optimal auction if the distribution is symmetric and max-ratio log-supermodular.*

In Appendix A we employ using the natural generalization of the max-ratio log-supermodularity condition to extend the above results to $m \geq 2$ items. The max-ratio representation of the density function f is a function f^{MR} satisfying

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

4.1 Max-ratio log-supermodular distributions

To understand MR-log-supermodular distributions better, we next show properties under which product distributions are MR-log-supermodular. We start proving that inequality (6) holds for MR-log-supermodular distributions. After that we will give a sufficient condition for an i.i.d. distribution to be MR-log-supermodular.

Lemma 15. *If f is a MR-log-supermodular function, then for any t_1 and θ,*

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

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

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

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

$$\frac{\mathrm{d}}{\mathrm{d}t_1}\left[\frac{\int_{y=0}^{t_1\theta} f(t_1, y)\ \mathrm{d}y}{\int_{y=0}^{t_1} f(t_1, y)\ \mathrm{d}y}\right] = 0. \tag{7}$$

□

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

$$f(z_1^{\lambda} z_2^{1-\lambda}) \geq f(z_1)^{\lambda} f(z_2)^{1-\lambda}, \qquad \forall \lambda \in [0,1], z_1, z_2.$$

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 16. *The product distribution on i.i.d. draws from a distirbution with geometric-geometric convex density is max-ratio log-supermodular.*

4.2 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 with finite outcome space can be converted to a unit-demand problem with dot-product utility function by letting $\boldsymbol{t}$ and $\boldsymbol{x}$ be vectors of size equal to the size of the outcome space, with t_i being the value for outcome i. For multi-item settings with m items, 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 17. *Bundle pricing is optimal if the max-ratio representation of the density function is MR-log-supermodular.*

5 Optimality of Bundle Pricing for Additive Preferences

In this section we provide sufficient 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 $\boldsymbol{t} \in T$ such that $t_1 \geq t_2$.

For $0 \leq s \leq 1$, define $F_{sum}(v) = \int_{\boldsymbol{t} \in T, t_1, t_2 \leq v} f \, \mathrm{d}\boldsymbol{t}$ to be the distribution of the agent's favorite item, and f_{sum} to be the corresponding density function.

Lemma 18 (single-dimensional projection)**.** *The expected revenue of a bundle pricing is equal to its expected virtual surplus with respect to vector field $\bar{\boldsymbol{\phi}}$ with $\bar{\phi}_1(\boldsymbol{t}) + \bar{\phi}_2(\boldsymbol{t}) = \phi_{sum}(t_1 + t_2)$ where $\phi_{sum}(v) = v - \frac{1 - F_{sum}(v)}{f_{sum}(v)}$ is the single-dimensional strong amortization for agent's sum of the values of items.*

Proof. Let $\boldsymbol{x}^p$ be the allocation corresponding to posting price p for the bundle, that is $x_1^p(\boldsymbol{t}) = x_2^p(\boldsymbol{t}) = 1$ if $t_1 + t_2 \geq p$, and $x_1^p(\boldsymbol{t}) = x_2^p(\boldsymbol{t}) = 0$ otherwise. We will show that the virtual surplus of $\boldsymbol{x}^p$ is equal to the revenue of posting price p, $R(p) = p(1 - F_{sum}(p))$. The virtual surplus is

$$\begin{aligned}\int_{\boldsymbol{t} \in T} (\boldsymbol{x}^p \cdot \boldsymbol{\phi} f)(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} &= \int_{\boldsymbol{t} \in T} \boldsymbol{x}^p(t_1, t_2) \cdot \bar{\boldsymbol{\phi}}(t_1, t_2) f(t_1, t_2) \, \mathrm{d}\boldsymbol{t} \\ &= \int_{\boldsymbol{t} \in T, t_1 + t_2 \geq p} \phi_{sum}(t_1 + t_2) f(t_1, t_2) \, \mathrm{d}\boldsymbol{t}.\end{aligned}$$

For a function h on T, define h^{SR} to be its transformation to sum-ratio coordinates, that is

$$h(t_1, t_2) = h^{SR}(t_1 + t_2, \frac{t_2}{t_1})$$

By performing a change of variables $s = t_1 + t_2$, and $\theta = t_2/t_1$, the virtual surplus of $\boldsymbol{x}^p$ can be written as

$$\begin{aligned}\int_{s=p}^{1}\int_{\theta=0}^{1}\phi_{sum}(s)f^{SR}(s,\theta)\frac{s}{1+\theta}\,\mathrm{d}\theta\mathrm{d}s &= \int_{s=p}^{1}\phi_{sum}(s)\int_{\theta=0}^{1}f^{SR}(s,\theta)\frac{s}{1+\theta}\,\mathrm{d}\theta\mathrm{d}s\\ &= \int_{s=p}^{1}\phi_{sum}(s)f_{sum}(s)\mathrm{d}s\end{aligned}$$

Replacing ϕ_{sum} by its definition,

$$\begin{aligned}&= \int_{s\geq p} t_1 f_{sum}(s) - (1-F_{sum}(s))\,\mathrm{d}s\\ &= -\int_{s\geq p}\frac{\mathrm{d}}{\mathrm{d}s}(s(1-F_{sum}(s))\,\mathrm{d}s\\ &= R(p) - R(1) = R(p).\end{aligned}$$

□

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

$$\begin{aligned}\bar{\phi}_1(\boldsymbol{t}) &= \frac{t_1}{t_1+t_2}\phi_{sum}(t_1+t_2) = t_1 - \frac{1-F_{sum}(t_1+t_2)}{f_{sum}(t_1+t_2)}\\ \bar{\phi}_2(\boldsymbol{t}) &= \frac{t_2}{t_1+t_2}\phi_{sum}(t_1+t_2) = t_2 - \frac{1-F_{sum}(t_1+t_2)}{f_{sum}(t_1+t_2)}.\end{aligned}$$

Notice that $\bar{\phi}_1$ and $\bar{\phi}_2$ have the same sign as ϕ_{sum}, 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 $\bar{\boldsymbol{\phi}}$, according to the following lemma.

Lemma 19 (consistency with virtual surplus maximization; incentive compatibility)**.** *Virtual surplus maximization according to vector field $\bar{\boldsymbol{\phi}}$ gives a bundle pricing p^* (and is incentive compatible) if and only if*

- $\bar{\phi}_1(\boldsymbol{t}), \bar{\phi}_2(\boldsymbol{t}) \geq 0$ *when* $t_1 + t_2 \geq p^*$ *and* $\bar{\phi}_1(\boldsymbol{t}), \bar{\phi}_2(\boldsymbol{t}) \leq 0$ *otherwise.*

Proof. We need to show that for the uniform price p, the allocation function $\boldsymbol{x}^p$ (defined in the proof of Lemma 8) optimizes $\boldsymbol{\phi}$ pointwise. Pointwise optimization of $\boldsymbol{x}\cdot\bar{\boldsymbol{\phi}}$ will result in $\boldsymbol{x} = (1,1)$ whenever $\bar{\phi}_1, \bar{\phi}_2 \geq 0$, and will result in $\boldsymbol{x} = (0,0)$ whenever $\phi_1, \phi_2 \leq 0$. □

In order to complete the construction of proof we should ensure that $\bar{\boldsymbol{\phi}}$ is an amortization of revenue. We already constructed $\bar{\boldsymbol{\phi}}$ such that the expected virtual surplus of the virtual surplus maximizer is equal to its revenue. So we only need to show that the virtual surplus of any incentive compatible allocation is at least its revenue.

Lemma 20. *The vector field $\bar{\boldsymbol{\phi}}$ is a weak amortization of revenue if for a strong amortization of revenue, and any incentive compatible allocation $\boldsymbol{x}$, and any s, θ,*

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

Proof. We invoke Lemma 7 and construct a strong amortization $\boldsymbol{\phi}$ that shows that the expected virtual surplus of any incentive compatible allocation is at least its revenue. Thus we need to show that the following expression is at lease zero, for all incentive compatible $\boldsymbol{x}$,

$$\int_{\boldsymbol{t}} \boldsymbol{x}(\boldsymbol{t}) \cdot [\bar{\boldsymbol{\phi}}(\boldsymbol{t}) - \boldsymbol{\phi}(\boldsymbol{t})] f(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} \tag{8}$$

We will now perform a change of variables $s = t_1 + t_2$, and $\theta = t_2/t_1$ on expression (8)

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

We can integrate by parts and write the above expression as

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

Define

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

which allows us to write

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

Under the assumption of the lemma, the expression inside the double integral is positive, which completes the proof. □

The following two lemmas show that the condition of Lemma 22 holds for all incentive compatible allocation under our regularity assumption (SR-log-supermodularity, to be defined shortly), completing the proof.

Lemma 21. *The allocation of any incentive compatible mechanism satisfies*

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

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

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

Letting ϵ approach zero implies the claim. □

So in order to complete the proof by combining Lemma 22 and Lemma 21, it is sufficient to show that

$$-\int_{y=\theta}^{1}[\bar{\phi}_1-\phi_1]^{SR}(s,y)f^{SR}(s,y)\frac{s}{1+y}\,\mathrm{d}y=\int_{y=\theta}^{1}[\bar{\phi}_2-\phi_2]^{SR}(s,y)f^{SR}(s,y)\frac{s}{1+y}\,\mathrm{d}y,$$

and that the two sides of the above equation are non-negative. We will show that by proving that a stronger condition holds in the following lemma. We say that a distribution is *sum-ratio* log-submodular (SR-log-submodular) if its sum-ratio representation is log-supermodular, that is,

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

Lemma 22. *If the distribution is SR-log-submodular, then there exist $\boldsymbol{\alpha}$ satisfying conditions of Lemma 5 such that $\gamma_1=-\gamma_2$, where $\gamma_1\geq 0$.*

We will here sketch the proof of the above lemma, and the complete proof appears in Appendix B. The requirement that $\gamma_1=-\gamma_2$ is equivalent to

$$\int_{y=\theta}^{1}[\boldsymbol{\phi}_1+\boldsymbol{\phi}_2-\bar{\boldsymbol{\phi}}_1-\bar{\boldsymbol{\phi}}_2]^{SR}(s,y)f^{SR}(s,y)\frac{s}{1+y}\,\mathrm{d}y=0.$$

For the above equation to hold for every s and θ, we must have

$$\boldsymbol{\phi}_1+\boldsymbol{\phi}_2=\bar{\boldsymbol{\phi}}_1+\bar{\boldsymbol{\phi}}_2=\lambda(t_1+t_2).$$

Given the fact that $\boldsymbol{\alpha}(\boldsymbol{t})=(\boldsymbol{t}-\boldsymbol{\phi}(\boldsymbol{t}))f(\boldsymbol{t})$,

$$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)(\boldsymbol{t})=(t_1+t_2-\lambda(t_1+t_2))f(\boldsymbol{t}).$$

The above equation, together with the requirement that $\boldsymbol{\alpha}$ is an amortization of utility (Lemma 5), will uniquely define $\boldsymbol{\alpha}$. Under SR-log-submodularity, the solution will satisfy $\gamma_1\geq 0$.

Theorem 23. *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 $(\bar{\boldsymbol{x}}^1(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n),\ldots,\bar{\boldsymbol{x}}^n(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^m))\in\mathcal{S}$ in which $\bar{\boldsymbol{x}}^\kappa(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)\in\mathbb{R}^m$ is the allocation of agent κ, together with a sequence of payment function $(p^1(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n),\ldots,p^n(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n))\in\mathbb{R}^n$ in which $p^\kappa(\boldsymbol{t}^1,\ldots,\boldsymbol{t}^n)\in\mathbb{R}$ 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,

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

where $\boldsymbol{t}^{-\kappa}$ 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,

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

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

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

where $\boldsymbol{t}^{-\kappa}$ 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 $(\boldsymbol{x}, p)$ defines a utility function $u(\boldsymbol{t}) = \boldsymbol{t} \cdot \boldsymbol{x}(\boldsymbol{t}) - p(\boldsymbol{t})$. Recall the following lemma from Section 2.3 which connects the utility function of an IC mechanism with its allocation function.

Lemma 3 (Rochet (1985))**.** *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 $\boldsymbol{x}(\boldsymbol{t}) = \nabla u(\boldsymbol{t})$.*

The payment function can be defined using the utility function and the allocation function as $p(\boldsymbol{t}) = \boldsymbol{t} \cdot \boldsymbol{x}(\boldsymbol{t}) - u(\boldsymbol{t})$. Using the above lemma, we can write payment to be $p(\boldsymbol{t}) = \boldsymbol{t} \cdot \nabla u(\boldsymbol{t}) - u(\boldsymbol{t})$. 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.

$$\begin{aligned} \max_{\bar{\boldsymbol{x}}^1, \ldots, \bar{\boldsymbol{x}}^n, u^1, \ldots, u^n} \quad & \sum_\kappa \int_{\boldsymbol{t}} [\boldsymbol{t} \cdot \nabla u^\kappa(\boldsymbol{t}) - u^\kappa(\boldsymbol{t})] f^\kappa(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} \\ & \forall \kappa, \boldsymbol{x}^\kappa = \nabla u^\kappa; u^\kappa \text{ is convex}, \\ & \forall \boldsymbol{t}, (\bar{\boldsymbol{x}}^1(\boldsymbol{t}^1, \ldots, \boldsymbol{t}^n), \ldots, \bar{\boldsymbol{x}}^n(\boldsymbol{t}^1, \ldots, \boldsymbol{t}^n)) \text{ is feasible allocation}. \end{aligned} \tag{9}$$

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

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

Now assume that for each agent κ, we have constructed $\boldsymbol{\alpha}^\kappa(\boldsymbol{t})$ according to Lemma 5. The above analysis and Program 9 tells us that we can write revenue as

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

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

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 $\boldsymbol{\phi}$ is a strong amortization of revenue by proving it satisfies conditions of Lemma 5, we now study its incentive compatibility in multi-agent settings. 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 $\{i|\sum_{j=m} x_{ij}=1\}\in\mathcal{S}$. A service-constrained environment is downward-closed if for any $S,S'\subseteq S$ such that $S\in\mathcal{S}$, we also have $S'\in\mathcal{S}$.

Notice that $\boldsymbol{\phi}_2\le\frac{t_2}{t_1}\boldsymbol{\phi}_1$ implies that when $\boldsymbol{\phi}_1\ge 0$, we have $\boldsymbol{\phi}_2\le\boldsymbol{\phi}_1$, and when $\boldsymbol{\phi}_1\le 0$, we also have $\boldsymbol{\phi}_2\le 0$. This implies that in service-constrained settings, an ex-post optimizer of $\boldsymbol{\phi}$ will assign positive probability only to the favorite item. In addition, if $\boldsymbol{\phi}(t)$ 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

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

which shows the convexity of the utility function, which in turn proves incentive compatibility (see Lemma 3). We have therefore proved the following lemma.

Lemma 24. *If $\boldsymbol{\phi}(\boldsymbol{t})$ is a non-decreasing function of t_1, and $\boldsymbol{\phi}_1\ge\boldsymbol{\phi}_2$ whenever $\boldsymbol{\phi}_2\ge 0$, then $\boldsymbol{\phi}$ is incentive compatible for any multi-agent downward-closed service-constrained environment.*

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 $\boldsymbol{\phi}_2\le\boldsymbol{\phi}_1$ when $\boldsymbol{\phi}_2\ge 0$. In particular, we can prove the following lemma (proof is in Appendix B).

Lemma 25. *There exists an ironed virtual value function $\bar{\phi}$ that is a weak amortization of revenue and is incentive compatible (by satisfying the conditions of Lemma 24) for MR-log-supermodular distributions.*

The above lemma directly implies our main theorem.

Theorem 26. *In multi-agent downward-closed service-constrained environments with MR-log-supermodular distributions, the optimal auction selects a feasible set maximizing the sum of virtual values for favorite items.*

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.

Giannakopoulos, Y. and Koutsoupias, E. (2014). Duality and optimality of auctions for uniform distributions. In *ACM Conference on Electronic Commerce*.

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

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.

McAfee, R. P. and McMillan, J. (1988). Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, 46(2):335–354.

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).

Pycia, M. (2006). Stochastic vs deterministic mechanisms in multidimensional screening.

Riley, J. and Zeckhauser, R. (1983). Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, 98(2):267–289.

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. (2014). Optimal mechanisms with simple menus. In *ACM Conference on Electronic Commerce*.

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 $T(t_1, \theta_2, \ldots, \theta_m) = \{\boldsymbol{t}' | t'_1 \geq t_1, t'_i \leq \theta_i t_1, \forall i \geq 2\}$. Define $C(t_1, \theta_2, \ldots, \theta_m) = \{\boldsymbol{t}' | t'_1 = t_1, t'_i \leq t'_i \theta_i, \forall i \geq 2\}$. Now define

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

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

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

The above definition implies that

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

As a result,

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

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

Finally, we verify that $\alpha_i(\boldsymbol{t}) \leq \frac{t_i}{t_1} \alpha_1(\boldsymbol{t})$. 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 $\boldsymbol{t}'$ 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 $\alpha_i(\boldsymbol{t}) \leq \frac{t_i}{t_1} \alpha_1(\boldsymbol{t})$ 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 $\alpha_i(\boldsymbol{t}) \leq \frac{t_i}{t_1} \alpha_1(\boldsymbol{t})$. 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 27. *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*

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

where

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

B Missing Proofs

B.1 Proof from Section 4

Proof of Lemma 8. Let $\boldsymbol{x}^p$ be the allocation corresponding to posting price p for item 1, that is $x_1^p(\boldsymbol{t}) = 1$ if $t_1 \geq p$, $x_1^p(\boldsymbol{t}) = 0$ otherwise, and $x_2^p(\boldsymbol{t}) = 0$ everywhere. We will show that the virtual surplus of $\boldsymbol{x}^p$ is equal to the revenue $p(1 - F_{\max}(p))$ of posting price p. The virtual surplus is

$$\begin{aligned}\int_{\boldsymbol{t} \in T} (\boldsymbol{x}^p \cdot \boldsymbol{\phi} f)(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} &= \int_{t_1 \geq 0} \int_{t_2 \leq t_1} \boldsymbol{x}^p(t_1, t_2) \cdot \boldsymbol{\phi}(t_1, t_2) f(t_1, t_2) \, \mathrm{d}t_2 \, \mathrm{d}t_1 \\ &= \int_{t_1 \geq p} \int_{t_2 \leq t_1} \phi_1(t_1, t_2) f(t_1, t_2) \, \mathrm{d}t_2 \, \mathrm{d}t_1.\end{aligned}$$

Since $\phi_1(t_1, t_2) = \phi_1(t_1)$, this becomes

$$\begin{aligned}&= \int_{t_1 \geq p} \phi_1(t_1) \int_{t_2 \leq t_1} f(t_1, t_2) \, \mathrm{d}t_2 \, \mathrm{d}t_1 \\ &= \int_{t_1 \geq p} \phi_1(t_1) f_{\max}(t_1) \, \mathrm{d}t_1\end{aligned}$$

Replacing ϕ_1 by its definition,

$$\begin{aligned}&= \int_{t_1 \geq p} t_1 f_{\max}(t_1) - (1 - F_{\max}(t_1)) \, \mathrm{d}t_1 \\ &= -\int_{t_1 \geq p} \frac{\mathrm{d}}{\mathrm{d}t_1} (t_1 (1 - F_{\max}(t_1)) \, \mathrm{d}t_1 \\ &= p(1 - F_{\max}(p)).\end{aligned}$$

□

Proof of Lemma 15. 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

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

□

Proof of Lemma 16. We need to prove

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

Recall that $f^{MR}(t_1, \theta_1) = f(t_1, t_1\theta_1)$. Since the distribution is a product one, this implies that $f^{MR}(t_1, \theta_1) = f_1(t_1) f_2(t_1\theta_1)$. 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

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

□

B.2 Proofs from Section 5

Proof of Lemma 22. We will construct $\boldsymbol{\alpha}$, and therefore indirectly γ, such that $\gamma_1 + \gamma_2 = 0$ and $\gamma_1 \geq 0$. By definition of γ we require

$$(\gamma_1 + \gamma_2)(s, \theta) = \int_{y=\theta}^{1} [\boldsymbol{\phi}_1 + \boldsymbol{\phi}_2 - \bar{\boldsymbol{\phi}}_1 - \bar{\boldsymbol{\phi}}_2]^{SR}(s, y) f^{SR}(s, y) \frac{s}{1+y} \, \mathrm{d}y = 0.$$

For the above equation to hold for every s and θ, we must have

$$\boldsymbol{\phi}_1 + \boldsymbol{\phi}_2 = \bar{\boldsymbol{\phi}}_1 + \bar{\boldsymbol{\phi}}_2 = \lambda(t_1 + t_2).$$

Given the fact that $\boldsymbol{\alpha}(\boldsymbol{t}) = (\boldsymbol{t} - \boldsymbol{\phi}(\boldsymbol{t}))f(\boldsymbol{t})$,

$$(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2)(\boldsymbol{t}) = (t_1 + t_2 - \lambda(t_1 + t_2))f(\boldsymbol{t}),$$

and therefore,

$$\frac{t_1}{t_1 + t_2}\lambda = t_1 - \frac{t_1}{t_1 + t_2}\frac{\alpha_1 + \alpha_2}{f}.$$

In order to have $\gamma_1 \geq 0$, it is sufficient to have $\boldsymbol{\phi}_1 - \bar{\boldsymbol{\phi}}_1 \geq 0$ everywhere,

$$\begin{aligned}
\boldsymbol{\phi}_1 - \bar{\boldsymbol{\phi}}_1 &= t_1 - \frac{\boldsymbol{\alpha}_1}{f} - \frac{t_1}{t_1 + t_2}\lambda \\
&= -\frac{\boldsymbol{\alpha}_1}{f} - \frac{t_1}{t_1 + t_2}\frac{\alpha_1 + \alpha_2}{f} \\
&= \frac{t_1}{(t_1 + t_2)f}(-\boldsymbol{\alpha}_1 \frac{t_2}{t_1} + \boldsymbol{\alpha}_2).
\end{aligned}$$

The following lemma shows that it is possible to construct $\boldsymbol{\alpha}$ to satisfy $-\boldsymbol{\alpha}_1 t_2/t_1 + \boldsymbol{\alpha}_2 \geq 0$. □

Lemma 28. *For SR-log-submodular distributions, there exists $\boldsymbol{\alpha}$ satisfying conditions of Lemma 5 such that $\boldsymbol{\phi}_1 + \boldsymbol{\phi}_2 = \lambda(t_1 + t_2)$, and $-\boldsymbol{\alpha}_1 t_2/t_1 + \boldsymbol{\alpha}_2 \geq 0$.*

Proof. The proof very closely resembles the proof of Lemma 12. The idea is that $(\alpha_1, \alpha_2) \cdot (-\theta, 1)$ is roughly (up to normalization) the derivative, with respect to s, of the following integral

$$\int_{C_1(s,\theta)} \boldsymbol{\alpha}(\boldsymbol{t}) \cdot \eta(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}.$$

By invoking the divergence theorem,

$$\int_{C_1(s,\theta)} \boldsymbol{\alpha} \cdot \eta = -\int_{T(s,\theta)} f - \int_{C_2(s,\theta)} \boldsymbol{\alpha} \cdot \eta$$

We differentiate the above equation with respect to s to get $(\alpha_1, \alpha_2) \cdot (-\theta, 1)$. In particular, for all s and θ we can write,

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

Now using the divergence theorem,

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

where the inequality holds by SR-log-submodularity of the distribution. Since f is always non-negative, this proves that $\alpha_2(t_1,t_2)\leq\frac{t_2}{t_1}\alpha_1(t_1,t_2)$.

□

B.3 Proofs from Section 6.3

Proof of Lemma 25. Since $\boldsymbol{\phi}_2(\boldsymbol{t})\leq\frac{t_2}{t_1}\boldsymbol{\phi}_1(\boldsymbol{t})$, the revenue of an allocation function can be upper bounded as follows,

$$\int\int\boldsymbol{x}(\boldsymbol{t})\cdot(\boldsymbol{\phi}_1(\boldsymbol{t}),\boldsymbol{\phi}_2(\boldsymbol{t}))f(\boldsymbol{t})\,\mathrm{d}(\boldsymbol{t})\leq\int\int\boldsymbol{x}(\boldsymbol{t})\cdot(1,\frac{t_2}{t_1})\boldsymbol{\phi}_1(\boldsymbol{t})f(\boldsymbol{t})\,\mathrm{d}\boldsymbol{t}$$

After a changing variables according to $\theta=\frac{t_2}{t_1}$, the upper bound becomes

$$\int_{t_1}\int_{\theta=0}^{1}\boldsymbol{x}^{MR}(t_1,\theta)\cdot(1,\theta)\boldsymbol{\phi}^{MR}(t_1,\theta)f^{MR}(t_1,\theta)t_1\,\mathrm{d}\theta\,\mathrm{d}t_1,$$

where $\boldsymbol{\phi}^{MR}(t_1,\theta)=\boldsymbol{\phi}(t_1,t_1\theta)$. For a function h on T, define h^{QR} to be its transformation to quantile-ratio coordinates, that is

$$h(t_1,t_2)=h^{QR}(\int_{T=T(t_1,1)}f(\boldsymbol{t})\,\mathrm{d}\boldsymbol{t},\frac{t_2}{t_1})$$

We perform another change of variables according to $q=\int_{T=T(t_1,1)}f(\boldsymbol{t})\,\mathrm{d}\boldsymbol{t}$, and rewrite the upper bound as

$$\int_q\int_\theta\boldsymbol{x}^{QR}(q,\theta)\cdot(1,\theta)\boldsymbol{\phi}^{QR}(q)\,\mathrm{d}\theta\,\mathrm{d}q.$$

Using integration by parts,

$$\int_q \int_\theta \boldsymbol{x}^{QR}(q,\theta)\cdot(1,\theta)\boldsymbol{\phi}^{QR}(q)\ \mathrm{d}\theta\ \mathrm{d}q = \int_q \int_\theta \frac{d}{dq}\boldsymbol{x}^{QR}(q,\theta)\cdot(1,\theta)R(q)\ \mathrm{d}\theta\ \mathrm{d}q$$

where $R(q)$ is the revenue of posting a price that is accepted with probability q. Now let $\bar{R}$ be the convex hull of function R. In particular, $\bar{R}'(q)$ is non-decreasing and $\bar{R}(q) \geq R(q)$. We next show that $\frac{d}{dq}\boldsymbol{x}^{QR}(q,\theta)\cdot(1,\theta) \geq 0$ for any incentive compatible allocation rule.

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

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

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

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

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

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

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

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

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

$$\int_q \int_\theta \frac{d}{dq}\boldsymbol{x}^{QR}(q,\theta)(1,\theta)R(q)\ \mathrm{d}\theta\ \mathrm{d}q \leq \int_q \int_\theta \frac{d}{dq}\boldsymbol{x}^{QR}(q,\theta)(1,\theta)\bar{R}(q)\ \mathrm{d}\theta\ \mathrm{d}q$$

We now study the allocation rule that optimizes virtual value function $(1,\theta)R'(q)$. Notice that when $\bar{R}(q) > R(q)$, we have that $\bar{R}'(q)$ is a constant and therefore $\frac{d}{dq}\boldsymbol{x}^{QR} = 0$. 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. □