Optimal Mechanisms with Simple Menus

Pingzhong Tang\textsuperscript{a}, Zihe Wang\textsuperscript{a,b}

\textsuperscript{a}Institute of Interdisciplinary Information Sciences
Tsinghua University, Beijing, China
\textsuperscript{b}Institute of Theoretical Computer Science
Shanghai University of Finance and Economics, Shanghai, China

Abstract

We consider revenue-optimal mechanism design for the case with one buyer and two items, when the buyer’s valuations are independent and additive. We obtain two sets of structural results of the optimal mechanisms, which can be summarized in one conclusion: under certain distributional conditions, the optimal mechanisms have simple menus.

The first set of results states that, under a condition that requires that the types are concentrated on lower values, the optimal menu can be sorted in ascending order. Applying the theorem, we derive a revenue-monotonicity theorem which states that stochastically dominated distributions yield less revenue.

The second set of results state that, under certain conditions which require that types are distributed more evenly or are concentrated on higher values, the optimal mechanisms have a few menu items. Our first result states that, for certain such distributions, the optimal menu contains at most 4 menu items. The condition admits power density functions. Our second result works for a weaker condition, under which the optimal menu contains at most 6 menu items. Our last result in this set works for the unit-demand setting, it states for uniform distributions, the optimal menu contains at most 5 items.

Keywords: Optimal mechanism design, menu representation

\hspace{1cm}Email addresses: kenshin@tsinghua.edu.cn (Pingzhong Tang), wzh588588@163.com (Zihe Wang)
1. Introduction

Revenue-optimal mechanism design has been a topic of intensive research over the past thirty years. The problem is, for a seller, to design a revenue-maximizing mechanism for selling $k$ items to $n$ buyers, given the buyers’ distributions of valuations but not the actual values themselves. A special case of the problem, where there is only one item ($k = 1$) and buyers have independent valuation distributions, has been resolved by Myerson’s seminal work (Myerson, 1981). Myerson’s approach has turned out to be very general and has been applied to a number of similar settings, such as (Maskin et al., 1989; Jehiel et al., 1996; Levin, 1997; Ledyard, 2007; Deng and Pekeč, 2011).

However, this line of work is limited because it does not deepen the understanding of the cases with more than one items ($k > 1$). In fact, even for the simplest multi-item case, where there are two independent items ($k = 2$) and one buyer ($n = 1$) with additive valuations, a direct characterization of the optimal mechanism is still open for general continuous valuation distributions.

For the special case of selling multiple, independent items to a single buyer, significant progress has been made in this particular setting lately. Hart and Nisan (2012) investigate the two simplest forms of auctions: selling the two items separately and selling them as a bundle. They prove that selling separately obtains at least one half of the optimal revenue while bundling always returns at least one half of the separate sale revenue. Hart and Nisan (2013) investigate how the “menu size” of an auction can affect the revenue and show that the revenue of any finite menu-sized auction can be arbitrarily far from the optimal (this implies that restricting attention to deterministic auctions, which have an finite-sized menu, indeed compromises generality). Carrol (2015) considers a robust version of the optimal mechanism design problem, where there is one buyer and multiple additive items and the seller only knows the marginal valuation distributions of each item but not the joint distribution. He shows that the worst-case (with respect to any joint distribution that is consistent with the marginal distributions) optimal mechanism is to separately sell each
item via a Myerson auction.

With respect to the literature of exactly optimal mechanism design, Thanassoulis (2004) provides examples where the optimal mechanism requires randomized allocations. Pycia (2006) further shows that in general, the optimal mechanism is randomized. (Manelli and Vincent, 2006, 2007; Pavlov, 2011a,b) obtain optimal mechanisms for several specific distributions (such as when both items are distributed according to the uniform [0,1] distribution). Daskalakis et al. (2013, 2015) study this problem from the perspective of duality theory. First they formulate the problem as a maximization problem over a convex domain and then consider its dual in the form of an optimal transportation problem. Their main result is a strong duality theorem, by applying the duality theorem, they can certify optimality by providing a complementary solution to the dual problem. Examples that illustrate their techniques include the optimality of an infinite-menu mechanism for two independent beta distributions, as well as optimality conditions for the grand bundling. Menicucci et al. (2015) prove sufficient conditions under which bundling is optimal for one buyer and two additive items. We will discuss the connection of this paper to our results in Section 5. Haghpanah and Hartline (2015) identify the sufficient and necessary conditions (include the independent uniform case) under which for one unit-demand buyer with two items, the optimal mechanism is to post a price for each item. We will discuss the connection of this paper to our result in the unit-demand section.

In the present paper, we study the case with one buyer and two independent items, in hope of a direct characterization of optimal mechanisms. We obtain several interesting structural results. Our conclusion is that, under some distributional conditions, optimal mechanisms have “simple” menus. We summarize our results into two parts, based on the conditions under which the results hold, as well as the different interpretations of “simplicity”.

For ease of presentation, we will use the following notation: for a density function $h$, the power rate of $h$ is $PR(h(x)) = \frac{h'(x)}{h(x)}$. 

3
Part I (Section 4). If the density functions $f_1$ and $f_2$ satisfy $\text{PR}(f_1(x)) + \text{PR}(f_2(y)) \leq -3$, $\forall x, y$, a condition that roughly states that the types are concentrated on lower values, the optimal mechanism has a monotone menu – sort the menu items in ascending order of payments, the allocation probabilities of both items increase simultaneously – a desirable property that fails to hold in general (cf (Hart and Reny, 2012)). Our result complements Hart and Reny’s observation and has two important implications.

1. (Hart and Nisan, 2012, Theorem 28). Hart and Nisan show that, if the two items are further identically distributed (i.e., $f_1 = f_2$), the bundling auction is optimal. Our result subsumes this theorem as a corollary.

2. A revenue monotonicity theorem: Based on the menu monotonicity theorem, we are able to prove that, stochastically dominated distributions yield less revenue, another desirable property that fails to hold in general.

Our proof is constructive and geometrical in the sense that we fix the buyer utility function on certain boundary lines of the valuation domain (according to the revenue formula, the seller’s revenue is not increasing in the buyer’s utility on these boundary lines, thus hard to analyze, so we fix this part of the utility function) and construct the remainder of the optimal utility function (for the remainder part of the valuation domain, the revenue is increasing in the buyer’s utility, according to the revenue formula). For several recent applications of the geometrical approach, see Wang and Tang (2015); Tang and Wang (2016); Tang et al. (2016).

Part II. (Section 5). If the density functions $f_1$ and $f_2$ satisfy $\text{PR}(f_1(x)) + \text{PR}(f_2(y)) \geq -3$ $\forall x, y$, a condition which roughly asserts that the types are distributed more evenly than the case described in Part I or are concentrated on higher values, the optimal mechanisms contain few menu items. In particular,
1. If both $PR(f_1(x))$ and $PR(f_2(y))$ are constants, the optimal mechanism contains at most 4 menu items. The result is tight. The constant power rate is satisfied by power functions $h(x) = ax^b$ and the uniform distribution as a special case. This is consistent with earlier results for uniform distributions (Manelli and Vincent, 2006; Pavlov, 2011a): the optimal mechanisms indeed contain four menu items.

2. – If $PR(f_1(x)) + PR(f_2(y)) = -3 \forall x, y$, the optimal mechanism contains at most 3 menu items.
– If $-2 \leq PR(f_1(x)) \leq y_A f_2(y_A) - 2$ and $-2 \leq PR(f_2(y)) \leq x_A f_1(x_A) - 2$, the optimal mechanism contains 3 menu items. Here $x_A$ and $y_A$ are the lowest possible valuations for item 1 and 2 respectively. Consequently, under either condition, selling the two items as a bundle yields at least half of the optimal revenue.

3. If we relax the condition to the case where $PR(f(x))$ is monotonically increasing, a fairly general condition satisfied by many distributions, the optimal mechanism is still simple in that it contains at most 6 menu items. This condition includes density functions such as exponential density and any density function whose Taylor series coefficients are nonnegative.

4. Our last result requires that the buyer demands at most one item. Under this condition, for uniform densities, the optimal mechanism contains at most 5 menu items. The result is also tight.

These results are in sharp contrast to Hart and Nisan’s recent result that there is some distribution where a finite number of menu items cannot guarantee any fraction of revenue (Hart and Nisan, 2013). Here we show that, for several wide classes of distributions, the optimal mechanisms have finite and simple menus. The conditions in these results are necessary; when the conditions do not hold, Daskalakis et al. (2013, 2015) show that, for a setting with two beta distributions, the optimal menu must consist
of a continuum of menu items.

Our proofs for this part are based on Pavlov’s characterization and geometrical analysis of how the revenue changes as a function of the utility of the buyer. The intuition is as follows: the “extreme points” in the set of convex utility functions on the boundary values are piecewise linear functions. By standard geometrical arguments, one can further show that these piecewise linear functions only contain a small number of pieces. Since the utility on inner values is linearly related to that on the boundary (because the gradient of the utility function on one direction must be 1 according to Pavlov (2011a,b)), it must be the case that the utility function on the inner points contains only a few linear pieces as well. In other words, the mechanism only contains a few menu items.

In the optimal auction design problem, bidders are utility maximizers. By incentive compatibility, the equilibrium utility as a function of the valuation must be convex. The hardness of optimal auction design is to maximize the seller’s revenue under the convex constraints. As one can expect, a common approach is to relax the convex constraint and compute the optimal solution of the relaxed problem. If one is lucky in that the relaxed optimal solution happens to be convex, an optimal solution is found. However this method fails sometimes. As mentioned, Daskalakis et al. (2015) transform the optimal mechanism design to the optimal transportation problem and give a procedure to certify the optimality of the auction. However, difficulties still exist when constructing the optimal solution to the transportation problem.

In parallel with this approach, we adjust the utility function while maintaining the convex constraints. We start from any convex utility, then try to increase or decrease the utility on every point and maintain the convex property in each small region.

Our results offer original insights of “what do optimal mechanisms look like?”, and are in line with the “simple versus optimal” literature (cf (Hartline and Roughgarden, 2009; Hart and Nisan, 2012)): in our case, simple mechanisms
are exactly optimal.

As for potential applications, the first set of results (Condition 1) apply to the scenario where the seller has a pair of second-hand items for sale and the buyer in general has low valuations. In this case, our analysis suggests that the seller should sell through a monotone menu. In particular, when the two items are symmetric, the seller should sell through bundling. Our second set of results apply to the scenario where the buyer is serious (Menicucci et al., 2015) in the sense that the buyer in general has high valuations, or the scenario where the buyer’s valuation is roughly uniform (e.g., the seller does not have much prior information about the buyer). In this case, our analysis suggests that it suffices for the seller to offer a small menu.

2. The setting

We consider a setting with one seller who has two distinct items for sale, and one buyer who has nonnegative private valuations $x$ for item 1, $y$ for the item 2, and $x + y$ for both items. The seller has zero valuation for items.

Both valuations $x$ and $y$ are unknown to the seller and are treated as independent random variables distributed according to density functions $f_1$ on $[x_A, x_B] \subset \mathbb{R}$ and $f_2$ on $[y_A, y_C] \subset \mathbb{R}$ respectively, where $x_A, y_A \geq 0$. The valuation (a.k.a. type) space of the buyer is then $V = [x_A, x_B] \times [y_A, y_C]$. To visualize, we sometimes refer to $V$ as rectangle $ABDC$, where $A$ represents the lowest possible type $(x_A, y_A)$ and $D$ represents the highest possible type $(x_B, y_C)$. Let $f(x, y) = f_1(x)f_2(y)$ be the joint density on $V$. We assume that both $f_1$ and $f_2$ are positive, bounded, and differentiable density functions.

The seller sells the items through a mechanism that consists of an allocation rule $q$ and a payment rule $t$. In our two-item setting, an allocation rule is represented by $q = (q_1, q_2)$, where $q_i$ is the probability that buyer gets item $i \in \{1, 2\}$. Given valuation $(x, y)$, the utility of the buyer is

$$u(x, y) = xq_1(x, y) + yq_2(x, y) - t(x, y)$$
In other words, buyer has a quasi-linear, additive utility function. It is sometimes convenient to view a mechanism as a (possibly infinite) set of menu items \( \{(q_1(x, y), q_2(x, y), t(x, y)) | (x, y) \in [x_A, x_B] \times [y_A, y_C]\} \). Given a mechanism, the expected revenue of the seller is \( R = \mathbb{E}_{(x, y)}[t(x, y)] \).

A mechanism must be Individually Rational (IR):

\[ \forall (x, y), u(x, y) \geq 0. \]

In other words, a buyer cannot obtain negative utility by participation.

By the revelation principle (Myerson (1981)), it is without loss of generality to focus on the set of mechanisms that are Incentive Compatible (IC):

\[ \forall (x, y), (x', y'), u(x, y) \geq xq_1(x', y') + yq_2(x', y') - t(x', y'). \]

This means that it is the buyer’s (weak) dominant strategy to report truthfully. Equivalently, an IC mechanism presents a set of menu items and asks the buyer to do the selection (a.k.a. the taxation principle (Vohra, 2011)). As a result,

\[ u(x, y) = \sup_{(x', y')} \{xq_1(x', y') + yq_2(x', y') - t(x', y')\}, \]

which is the supremum of a set of linear functions of \((x, y)\). Thus, \( u \) must be convex. Fixing \( y \), by IC, we have

\[
\begin{align*}
u(x', y) - u(x, y) &- q_1(x, y)(x' - x) \\
&= x'q_1(x', y) + yq_2(x', y) - t(x', y) - xq_1(x, y) - yq_2(x, y) \\
&+ t(x, y) - x'q_1(x, y) + xq_1(x, y) \\
&= x'q_1(x', y) + yq_2(x', y) - t(x', y) - (x'q_1(x, y) + yq_2(x, y) - t(x, y)) \geq 0
\end{align*}
\]

Substitute \( x' \) twice by \( x^- = x - \epsilon \) and \( x^+ = x + \epsilon \) respectively. For any arbitrarily small positive \( \epsilon \), we have

\[
u_x(x^-, y) \leq q_1(x, y) \leq u_x(x^+, y),
\]

where \( u_x \) denotes the partial derivative of \( u \) on the \( x \) direction. The inequality above implies that \( u \) is differentiable almost everywhere on \( x \) and \( u_x = q_1(x, y) \).
Similarly, \( u \) is differentiable almost everywhere on \( y \) and \( u_y = q_2(x, y) \). As a result, \( u_x \) and \( u_y \) must be within the interval \([0, 1]\). This means that the seller can never allocate more than one unit of either item. The payment function \( t \) can now be represented in terms of the utility function \( u \). If \( u \) is differentiable at point \((x, y)\), we have 
\[
t(x, y) = xu_x(x, y) + yu_y(x, y) - u(x, y). \tag{1}
\]

The seller’s problem is to design a non-negative, convex utility function, whose partial derivative on either \( x \) or \( y \) must be within \([0, 1]\), that maximizes the expected revenue \( R \) (cf. (Hart and Nisan, 2012, Lemma 5)).

### 3. Representing revenue as a function of the buyer’s utility

Let \( \partial \Omega \) be a positively oriented (i.e., the path integral is in the counter-clockwise direction of its boundary), piecewise smooth, simple, closed curve and let \( \Omega \) be the region enclosed by the curve \( \partial \Omega \). Let \( z = (x, y)^T \) and \( \mathcal{J}(z) = zu(z)f(z) \). Under this condition, we can apply Green’s Theorem. We get
\[
\int_{\Omega} \nabla \cdot \mathcal{J} \, dz = \oint_{\partial \Omega} \mathcal{J} \cdot \hat{n} \, ds,
\]
where \( \hat{n} \) is the outward-pointing unit normal vector on the boundary. Furthermore,
\[
\begin{align*}
\nabla \cdot \mathcal{J} &= 2u(z)f(z) + (\nabla u(z))^T zf(z) + u(z)z^T \nabla f(z) \\
&= [(\nabla u(z))^T z - u(z)]f(z) + [3f(z) + z^T \nabla f(z)]u(z) \\
&= t(z)f(z) + \triangle(z)u(z)
\end{align*}
\]

where \( \triangle(x, y) = 3f_1(x)f_2(y) + xf'_1(x)f_2(y) + yf'_2(y)f_1(x) \).

The seller’s revenue formula within \( \Omega \) is as follows:
\[
R_{\Omega} = \int_{\Omega} t(z)f(z) \, dz = \int_{\Omega} (\nabla \cdot \mathcal{J} - \triangle(z)u(z)) \, dz
\]
\[
= \int_{\partial \Omega} \mathcal{J} \cdot \hat{n} \, ds - \int_{\Omega} \triangle(z)u(z) \, dz.
\]

\(^1\)If \( u \) is not differentiable at type \((x, y)\), we let the payment be the largest payment in the small neighborhood of the current type, i.e., \( t(x, y) = \lim_{\epsilon \to 0} \sup\{t(\tilde{x}, \tilde{y}) : u \text{ is differentiable at type } (x, y) \text{ and } |\tilde{x} - x| + |\tilde{y} - y| < \epsilon\} \). In fact, the payments on such types do not affect the seller’s revenue, since it can be shown that the number of such points is countable.
Now let \( \Omega \) to be the rectangle \( ABDC \), the seller’s total revenue \( R_{ABDC} \) is

\[
\int_{y_A}^{y_C} x_B u(x_B, y) f_1(x_B) f_2(y) \, dy + \int_{x_A}^{x_B} y_C u(x, y_C) f_1(x) f_2(y_C) \, dx
- \int_{y_A}^{y_C} x_A u(x_A, y) f_1(x_A) f_2(y) \, dy - \int_{x_A}^{x_B} y_A u(x, y_A) f_1(x) f_2(y_A) \, dx
- \int_{x_A}^{x_B} \int_{y_A}^{y_C} u(x, y) \triangle(x, y) \, dy \, dx
\]  

(1)

Formula (1) consists of five terms. The first term represents the part of the seller’s revenue that depends on the utilities on edge \( BD \) only. Moreover, this part is increasing as the utilities on edge \( BD \) increase. Similarly, the second term represents the part of the seller’s revenue that depends positively on the utilities on edge \( CD \). The third and fourth terms represent respectively the parts of the seller’s revenue that depend negatively on the utilities on edges \( AC \) and \( AB \). The fifth term represents the part of the revenue that depends on the utilities on the inner points of the rectangle. Under different conditions, \( \triangle(x, y) \) can be either positive or negative, which suggests this part can either increase or decrease as the utilities on inner points increase. We now explicitly define these conditions.

**Definition 3.1.** For density \( h(x) \), let \( PR(h(x)) = \frac{xh'(x)}{h(x)} \) be its power rate.

Power rate is also known as the elasticity of density, since \( \frac{h'(x) \, dx}{h(x)} \rightleftarrows \frac{xh'(x)}{h(x)} \).

Consider the following two conditions regarding the power rate.

**Condition 1:** \( PR(f_1(x)) + PR(f_2(y)) \leq -3, \forall(x, y) \in V. \)

**Condition 2:** \( PR(f_1(x)) + PR(f_2(y)) \geq -3, \forall(x, y) \in V. \)

Under Condition 1, we have \( \triangle(x, y) \leq 0 \). This means that the seller’s revenue depends positively on the utilities of the inner points. Similarly, under Condition 2, the seller’s revenue depends negatively on the utilities of the inner points.

**Remark 1.** To understand the intuition behind the power rate, consider an example of selling only one item, where the valuation distribution is uniform on
[0, 1] (in this example, we have a relatively high power rate: \( PR = 0 \)). Consider a mechanism that has 2 menu items: \((0, 0)\) and \((1, 0.5)\) (take-it-or-leave-it on price 0.5). Now let us consider the effect of adding a new menu item \((0.5, 0.2)\).

- **Case 1.** When the buyer’s valuation is within \([0.4, 0.5]\), the utility of the buyer weakly increases (compared to the old mechanism) by switching to this new menu item. The seller’s revenue also increases because of positive probability of sale.

- **Case 2.** When the buyer’s valuation is within \([0.5, 0.6]\), the utility of the buyer also weakly increases by switching to this new menu item. However, the seller’s revenue decreases because the buyer now chooses a lower payment menu item.

Intuitively, a high power rate \((PR \geq -3)\) places sufficiently high density on large valuations, which ensures that the revenue increment in Case 1 is less than the revenue decrement in Case 2. Therefore, adding more menu items hurts revenue, if such menu items only change the utility of the inner points. In other words, under Condition 2 (the high power rate case), in order to maximize revenue, one must only keep the menu items that are chosen by the types on the boundaries. This leads to the conclusion that the number of menu items in the optimal mechanism must be small.

Based on the two conditions above, we obtain two sets of results: under Condition 1, the optimal mechanisms have simple menus in the sense that their menus are monotonically increasing – allocation probabilities and payments are increasing in the same order. Under Condition 2, the optimal mechanisms also have simple menus, but in a different sense, i.e. that their menus contain few items.

4. **Part I: menu monotonicity and revenue monotonicity**

In this section, we consider the case where the power rates of both density functions satisfy Condition 1. When this condition is not met, Hart and Reny
(2012) give several interesting counter-examples of revenue monotonicity: the optimal revenue for stochastically dominated valuation distributions may be greater than that of stochastically dominating distributions. When this condition is met, for identical item distributions, Hart and Nisan (2012) prove that bundling is optimal.

In this section, we show that, under Condition 1, the optimal menu can be sorted so that both allocations as well as payments monotonically increase. We coin this result the menu monotonicity theorem. The theorem has two corollaries. First, it yields a version of the revenue monotonicity theorem that complements the Hart-Reny result above. Second, it subsumes the above Hart-Nisan result.

Our analysis starts from a simple observation: any optimal mechanism must extract all the buyer’s rent when he has the lowest type.

**Lemma 4.1.** In the optimal mechanism, \( u(x_A, y_A) = 0 \).

**Proof.** Suppose otherwise that \( u(x_A, y_A) > 0 \), one can revise every menu item from \( (q_1(x, y), q_2(x, y), t(x, y)) \) to \( (q_1(x, y), q_2(x, y), t(x, y) + u(x_A, y_A)) \) and obtain a mechanism with strictly higher revenue, contradiction. \( \square \)

**Theorem 4.2. Menu Monotonicity**

Under Condition 1, the menu items of the optimal mechanism can be represented in the form \( (q_1(t), q_2(t), t) \), such that allocation probabilities \( q_1(t) \) and \( q_2(t) \) are weakly increasing in the payment \( t \).

Roughly speaking, Theorem 4.2 suggests that, among the menu items of the optimal mechanism, a higher payment \( t \) corresponds to higher allocation probabilities \( q_1 \) and \( q_2 \). Note that allocation and payment monotonicity are well understood in the single-item optimal auction (i.e., the Myerson auction) but in general fail to hold in two-item settings (Hart and Reny, 2012).

In the following, we give a constructive proof. By Formula (1), under Condition 1, we know that the seller’s revenue increases as the utility of the buyer increases on \( V \), except on edges \( AB \) and \( AC \). Our idea is to fix the utility
function on $AB$ and $AC$ and construct the (largest possible) remainder of the utility function subject to convexity.

Proof. Look at Figure 1, we use $u(AB)$ to denote the buyer utility function on edge $AB$; others notations, such as $u(AC)$, are similar.

Fixing any $u(AB)$ and $u(AC)$ (not necessarily optimal), consider any point $(x_A, y_A, z)$ lower than $A(x_A, y_A, 0)$ on the vertical line $(x = x_A, y = y_A)$, draw a plane going through the point and touching $u(AC)$ and $u(AB)$ (subject to that the gradient of the plane on each direction is no greater than 1). When $|z|$ is sufficiently high, it is possible that the plane cannot touch $u(AC)$ even if the gradient is set to be 1 in the $y$ direction. In this case, we define the gradient of the plane in the $y$ direction to be 1, without touching $u(AC)$. We treat the case where the plane cannot touch $u(AB)$ even if the gradient on the $x$ direction is set to be 1 similarly.

It is easy to see that the definition above defines a unique plane for each point $(x_A, y_A, z)$. We call this plane $u^*$.

We now claim that $u^*(x, y) = \sup_{z \in (-\infty, 0]} \{u^z(x, y)\}$ is the optimal utility function subject to fixed $u(AB)$ and $u(AC)$. To visualize, one can imagine walking down the vertical line passing through $A$, while simultaneously shooting a plane touching $u(AB)$ and $u(AC)$. The optimal utility function is simply the
supremum of all such planes. Now we prove this claim.

Since \( u^* \) is the supremum of a set of planes, \( u^* \) must be convex.

First, we prove \( (x, y) \in AB \cup AC \), \( u^*(x, y) = u(x, y) \). Pick any point \( (x_0, y_0) \in AB \). We have \( u^*(x_0, y_0) \leq u(x_0, y_0) \) for \( z \in (-\infty, 0] \), so \( u^*(x_0, y_0) = \sup_{z \in (-\infty, 0]} \{ u^z(x_0, y_0) \} \leq u(x_0, y_0) \). Since \( u(AB) \) is convex, there always exists a plane \( u^{z_0} \), where \( z_0 = u(x_0, y_0) - q_1(x_0, y_0)(x_0 - x_A) \), that passes through \( (x_0, y_0, u(x_0, y_0)) \). So \( u^*(x_0, y_0) \geq u^{z_0}(x_0, y_0) = u(x_0, y_0) \), i.e. \( u^*(x_0, y_0) = u(x_0, y_0) \). Similar for points on \( AC \).

Second, we prove \( (x, y) \in V \setminus \{ AB \cup AC \} \), \( u^*(x, y) \) is the largest possible value subject to fixed \( u(AB) \) and \( u(AC) \). Pick any point \( (x_1, y_1) \in V \setminus \{ AB \cup AC \} \). Let the largest possible utility on point \( (x_1, y_1) \) be \( \tilde{u}(x_1, y_1) \), achieved by utility function \( \tilde{u} \). Let \( (x_1, y_1, \tilde{u}(x_1, y_1)) \) be in some plane \( \tilde{u}^{(x_1, y_1)}(x, y) = x \tilde{q}_1(x_1, y_1) + y \tilde{q}_2(x_1, y_1) - \tilde{T}(x_1, y_1) \).

In other words, \( \tilde{u}^{(x_1, y_1)}(x, y) \) is the utility of the buyer at type \( (x, y) \) but chooses the menu item \( (\tilde{q}_1(x_1, y_1), \tilde{q}_2(x_1, y_1), \tilde{T}(x_1, y_1)) \). By IC, \( \tilde{u}^{(x_1, y_1)}(x, y) \leq \tilde{u}(x, y) \). Think of \( \tilde{u}^{(x_1, y_1)}(x, y) \) as a plane that is always weakly below \( \tilde{u}(x, y) \) but touches \( \tilde{u} \) at the point of \( (x_1, y_1) \).

Plane \( \tilde{u}^{(x_1, y_1)} \) passes through point \( (x_A, y_A, z_1) \), where \( z_1 = x_A \tilde{q}_1(x_1, y_1) + y_A \tilde{q}_2(x_1, y_1) - \tilde{T}(x_1, y_1) \). By definition, \( u^{z_1} \) goes through point \( (x_A, y_A, z_1) \) too. By our construction, \( u^{z_1} \) has the largest possible gradients in both directions subject to the fixed \( u(AB) \) and \( u(AC) \). Since we fix \( u(AB) \) and \( u(AC) \),

\[ \tilde{u}(AB) = u(AB), \quad \tilde{u}(AC) = u(AC), \]  

so \( u^{z_1} \) has weakly larger gradients than \( \tilde{u}^{(x_1, y_1)} \). Hence, we have \( u^*(x_1, y_1) \geq u^{z_1}(x_1, y_1) \geq \tilde{u}^{(x_1, y_1)}(x_1, y_1) \). Since \( (x_1, y_1) \) is arbitrarily chosen and \( \tilde{u}^{(x_1, y_1)} \) is the largest at \( (x_1, y_1) \), we conclude that \( u^*(x, y) \) is the largest possible value on any point \( (x, y) \) subject to fixed \( u(AB) \) and \( u(AC) \).

Finally, according to Formula (1), \( u^* \) gives the optimal revenue subject to fixed \( u(AB) \) and \( u(AC) \). We now prove that our construction of \( u^* \) is an upper envelope of a set of monotonically increasing planes \( u^z \).

Let \( t^z \) denote the payment in the menu item according to the plane \( u^z \). For any valuation \( (x, y) \), we have \( u^z(x, y) = xu^*_x(x, y) + yu^*_y(x, y) - t^z \). Consider the
valuation \((x_A, y_A)\), we have
\[
t^z = x_A u^z_A(x_A, y_A) + y_A u^z_B(x_A, y_A) - z
\]

In the figure, the payment is the intercept between the origin and the cross point of the plane \(u^z\) and \(z\)-axis. When \(|z|\) increases, the cross point descends, the gradients of the plane in both directions increase and the payment increases. In other words, allocation probabilities weakly increase and payment strictly increases. This completes the proof. \(\square\)

**Remark 2.** This theorem can be extended to the case where the valuations of the two items are correlated. In this case, let \(\Delta(x, y) = 3f(x, y) + x f'_x(x, y) + y f'_y(x, y)\). Condition 1 becomes \(\frac{x f'_x(x, y) + y f'_y(x, y)}{f(x, y)} \leq -3\). The proof of the theorem remains the same.

Theorem 4.2 implies the aforementioned Hart-Nisan result as a corollary.

**Corollary 4.3.** *(Hart and Nisan, 2012, Theorem 28)* For 2 i.i.d. items, \(PR(f_1) = PR(f_2) \leq -\frac{3}{2}\), bundling is optimal.

**Proof.** For i.i.d distributions, it is without loss of generality to restrict our attention to mechanisms that are symmetric with respect to the items (Maskin and Riley, 1984). This means that for any pair of buyer’s valuations \((x, y)\) and \((y, x)\), the utilities are the same, i.e., \(u(x, y) = u(y, x)\). So, \(u(AB)\) is symmetric to \(u(AC)\). As a result, any plane \(u^z\) constructed in the proof of Theorem 4.2 is symmetric in the sense that its gradient on both directions is the same. In other words, \(q_1(v) = q_2(v) \forall v \in V\). That is, the two items must be sold with the same probability. So the revenue maximization problem subject to this constraint can be thought of as a posted pricing problem for the bundle. From the theory of single-item auctions, the optimal posted pricing must be deterministic, i.e., a posted price for the bundle. \(\square\)

Menicucci et al. (2015) derive sufficient conditions under which bundling is optimal. The case they study is Condition 2 listed in this paper, and therefore it does not overlap with the result above, which holds under Condition 1.
As another application of Theorem 4.2, we obtain a revenue monotonicity theorem in this setting.

**Theorem 4.4. (Revenue Monotonicity)**

*Under Condition 1, the optimal revenue is increasing:* let $F_i, G_i$ be the cumulative distribution function of density functions $f_i, g_i, i = 1, 2$, respectively. If $G_1$ and $G_2$ first-order stochastically dominate $F_1$ and $F_2$, respectively, the optimal revenue obtained for $(G_1, G_2)$ is no less than that of $(F_1, F_2)$.

Intuitively, revenue monotonicity is desirable in the sense that, when a seller puts effort (e.g., advertisement campaigns) into improving the buyer’s valuations, say from $(F_1, F_2)$ to $(G_1, G_2)$, he is guaranteed to obtain more revenue, if he commits to using optimal mechanisms. Additional discussion on menu and revenue monotonicity can be found in (Hart and Reny, 2012).

**Proof.** Consider any two points $(x_2, y_2)$ and $(x_3, y_2)$, where $x_3 > x_2$. If $q_1(x_2, y_2) < q_1(x_3, y_2)$, by Theorem 4.2, we must have $t(x_2, y_2) < t(x_3, y_2)$. If $q_1(x_2, y_2) = q_1(x_3, y_2)$, then $q_1(x, y_2) = q_1(x_2, y_2), \forall x \in [x_2, x_3]$. $u(x_3, y_2) = u(x_2, y_2) + q_1(x_2, y_2)(x_3 - x_2)$, which can be achieved by choosing the same menu item as the one chosen by $(x_2, y_2)$. While the buyer at type $(x_3, y_2)$ has several menu items that all achieve the highest utility, we can assume, WLOG, that the buyer chooses the menu item with the highest payment ((Hart and Reny, 2012)). Thus there is an optimal choice guarantees $t(x_2, y_2) \leq t(x_3, y_2)$.

To sum up, $t(x_2, y_2) \leq t(x_3, y_2)$ when $x_2 \leq x_3$. For the same reason, $t(x_3, y_2) \leq t(x_3, y_3)$ when $y_2 \leq y_3$. Hence $t(x, y)$ is a weakly monotone increasing function in both directions. Suppose $G_1$ and $G_2$ first-order stochastically dominates $F_1$ and $F_2$ respectively. Let $R(F_1 \times F_2)$ denote the optimal revenue when item 1 and 2 distribute independently according to $F_1$ and $F_2$. For the distribution $G_1 \times G_2$, we adopt the mechanism which is optimal for the distribution $F_1 \times F_2$, and denote the revenue by $R^*(G_1 \times G_2)$. We

---

$^2G_i$ first-order stochastically dominates $F_i$ if $G_i(x) \leq F_i(x)$ for all $x$ and $G_i(x) < F_i(x)$ for some $x$. 

16
have $R^*(G_1 \times G_2) \geq R(F_1 \times F_2)$, since $t$ is weakly increasing. By transitivity, \\
$R(G_1 \times G_2) \geq R^*(G_1 \times G_2) \geq R(F_1 \times F_2)$. □

5. Part II: Optimal mechanisms with small menus

In this section, we investigate optimal mechanisms under Condition 2. We obtain several results saying that the optimal mechanism contains only a few menu items. All these results are built upon Pavlov’s characterization (Pavlov, 2011a) and an important lemma introduced in the next subsection.

5.1. Pavlov’s characterization and the graph representation lemma

If both $f_1$ and $f_2$ satisfy Condition 2, Pavlov (2011a, Proposition 2) states that, in the optimal mechanism, the seller either keeps both items (i.e., $q = (0,0)$), or sells one of the items at probability 1 (i.e., $q_1 = 1$ or $q_2 = 1$).

For a graphical representation, let the buyer’s valuation be within the rectangle $ABDC$. We have the following lemma.

**Lemma 5.1. Graph Representation Lemma**

Under Condition 2, the optimal mechanism can either be in the form of the rectangle shown in Fig. 2 or the one shown in Fig. 3. Formally speaking, the optimal mechanism divides the valuation rectangle into four regions,

1. in the bottom-left region (region ASME in both figures), it assigns $q = (0,0)$ and $u(x,y) = 0$ to any point $(x,y)$ in the region. Furthermore, region ASME is convex.
2. in the top-right region, it assigns $q = (1,1)$ to any point in the region.
3. in the top-left region, it assigns $q = (\ast,1)$ to any point in the region, where $\ast$ is a variable. The region consists of a set of vertical line segments, each corresponds to a distinct menu item.

Some regions may be empty. It is also possible that “zero region” crosses $CD$ or $BD$. Our techniques and results still apply to that case.
4. Symmetrically, in the bottom-right region, it assigns \( q = (1, *) \) to any point in the region. The region consists of a set of horizontal line segments, each corresponds to a distinct menu item.

5. The boundary between the top-left and top-right regions is vertical (QL in both figures); the boundary of the top-right and bottom-right regions is horizontal (MN in Fig. 2 or LI in Fig. 3). The curve ML in Fig. 3 is monotonically increasing.

![Figure 2](image)

**Figure 2:** The optimal allocation that there is a point on curve SME choosing the menu item (1,1).

![Figure 3](image)

**Figure 3:** The optimal allocation that there is no point on curve SME choosing the menu item (1,1).

**Proof.** We first determine the relative positions of the four possible regions.

If the seller keeps both items, the utility of the buyer is zero. Since \( u(x, y) \) is an increasing function, it assigns \( q = (0, 0) \) in the bottom-left region, i.e. ASME. Since \( u(x, y) \) is convex, the convex combination of any two zero-utility points must also be zero. Therefore, ASME is a convex region.

If for a type \((x_0, y_0)\) with \( q_1(x_0, y_0) = q_2(x_0, y_0) = 1 \), for any point \((x_1, y_1)\), IC requires that

\[
\begin{align*}
    x_0 + y_0 - t(x_0, y_0) &\geq x_0 q_1(x_1, y_1) + y_0 q_2(x_1, y_1) - t(x_1, y_1), \\
    x_1 q_1(x_1, y_1) + y_1 q_2(x_1, y_1) - t(x_1, y_1) &\geq x_1 + y_1 - t(x_0, y_0).
\end{align*}
\]

Summing the two inequalities, we get \((q_1(x_1, y_1) - 1)(x_1 - x_0) + (q_2(x_1, y_1) - 1)(y_1 - y_0) \geq 0\). If \(x_1 > x_0, y_1 > y_0\), we must have \(q_1(x_1, y_1) = q_2(x_1, y_1) = 1\).
Let \((x_2, y_2)\) be a point where some positive proportions of the items are sold, then according to Pavlov’s characterization (Pavlov, 2011a), one of the items must be sold deterministically. Consider two types \((x_2, y_2)\) and \((x_3, y_3)\) where 

\[ q_1(x_2, y_2) = 1, \quad q_2(x_2, y_2) < 1 \quad \text{and} \quad q_1(x_3, y_3) < 1, \quad q_2(x_3, y_3) = 1. \]

By IC, we must have

\[ x_2 + y_2q_2(x_2, y_2) - t(x_2, y_2) \geq x_2q_1(x_3, y_3) + y_2 - t(x_3, y_3), \]

\[ x_3q_1(x_3, y_3) + y_3 - t(x_3, y_3) \geq x_3 + y_3q_2(x_2, y_2) - t(x_2, y_2). \]

Summing up the two inequalities, we get 

\[ (1 - q_1(x_3, y_3))(x_2 - x_3) + (1 - q_2(x_2, y_2))(y_3 - y_2) \geq 0. \]

So, one of \(x_2 < x_3\) and \(y_2 > y_3\) does not hold. This implies the second part of (5).

To sum up, \((1, 1)\) must be assigned to the upper right corner, \((1, q_2(x, y))\) is assigned to the bottom-right corner, \((q_1(x, y), 1)\) is assigned to the upper left corner, and \((0, 0)\) is assigned to the bottom-left corner, (some regions may be empty).

Let the allocation vector at \((x_4, y_4)\) be \((1, q_2(x_4, y_4))\). For any \(x \in [x_4, x_B]\), by IC, we must have \(q_1(x, y_4) = 1\), and

\[ u(x, y_4) = u(x_4, y_4) + x - x_4 = x_4 + y_4 \cdot q_2(x_4, y_4) + x - x_4 = x + y_4 \cdot q_2(x_4, y_4) \]

In other words, choosing menu item \((1, q_2(x_4, y_4))\) is one of the best responses for buyer at \((x, y_4)\). This implies the first part of (5): the boundary between for different \(q_2\) in \((1, q_2(x, y))\) is horizontal. In particular, in Fig. 2, \(MN\) is horizontal. Similarly, \(LQ\) is vertical.

If there is a point on curve \(SE\) choosing the menu item \((1, 1)\), the mechanism is of the form shown in Fig. 2, otherwise it is of the form shown in Fig. 3. □

The mechanism based on Fig. 2 is a mechanism of a zero set defined in Daskalakis et al. (2015). The zero set is the region \(ASLME\) that has zero utility. The utility function satisfies the condition:

\[ u(x) = \min_{z \in ASLME} \|z - x\|_1 \]
The closed form of this kind for the optimal mechanism can be calculated by the procedure shown in subsection 7.2.2 in Daskalakis et al. (2015). However, we do not know any optimal mechanism that is of the form shown in Fig. 3.

5.2. Optimal mechanisms for constant power rate

To describe our first theorem under Condition 2, we need the following condition on density functions.

**Condition 3:** \( PR(f_i(x)) \), \( i = 1, 2 \), is constant.

**Theorem 5.2.** Under Conditions 2 and 3, there is an optimal mechanism that has at most 4 menu items.

The result is tight: one can find instances where the optimal mechanism contains exactly 4 menu items (Pavlov, 2011a, Example 3).

![Figure 4: The optimal allocation that there is no point on curve SME choosing (1,1) allocation menu item.](image)

We prove the theorem for the case shown in Fig. 4. (All figures except Fig. 2 and 3 are used as intermediate figures to illustrate the proofs, not the final shapes of the optimal mechanisms. For example, in this proof, we start from an arbitrary diagram that only has the properties listed in Lemma 5.1.) The other case related to Fig. 2 follows from an almost identical proof. In Fig. 4, there are two sub-cases to be considered. We first prove two lemmas for the case when there is no point on CA choosing menu item \((1,\ast)\). We will discuss
the case when there is some point on $CA$ choosing menu item $(1, \ast)$ in the proof of Theorem 5.2.

We draw a horizontal line through $M$ which intersects $BD$ at $N$. Then, we draw a vertical line through $M$ crossing $CD$ at $G$. We have the following two lemmas.

**Lemma 5.3.** There exists an optimal utility function such that $u(BN)$ is piecewise linear with at most 2 pieces, when there is no point on $CA$ choosing menu item $(1, \ast)$.

![Figure 5: The optimal allocation that there is no point on curve SME choosing allocation menu item (1,1).](image)

**Proof.** For any point $K$ on $BN$ (see Fig.5), draw a horizontal line that intersects curve $SE$ at $W$. Because $u_x = 1$ for all the types in region $EBNM$, we can represent the utility of any point $(x,y_K)$ on line $KW$ using $u(K)$. Formally, $u(x,y_K) = u(K) + x - x_K, \forall x \in [x_W,x_K]$. The revenue obtained in $EBNM$ is

$$R_{EBNM} = \int_{EBNM} \mathcal{J} \cdot \hat{\mathbf{n}} ds - \int_{EBNM} \Delta(z)u(z)dz$$

$$= \int_{BN} \mathcal{J} \cdot \hat{\mathbf{n}} ds - \int_{y_A}^{y_B} \int_{x_B-u(x_B,y)}^{x_B} \Delta(x,y)u(x,y)dxdy + \int_{NMEB} \mathcal{J} \cdot \hat{\mathbf{n}} ds$$

$$= \int_{y_A}^{y_B} \left[ x_Bu(x_B,y)f_1(x_B)f_2(y) \right. \right.$$  

$$- \left. \int_{x_B-u(x_B,y)}^{x_B} \Delta(x_B,y)(u(x_B,y) + x - x_B)dx \right] dy + c$$
Let \( R(u(x_B, y), y) = x_B u(x_B, y) f_1(x_B) f_2(y) - \int_{x_B - u(x_B, y)}^{x_B} \Delta(x_B, y)(u(x_B, y) + x - x_B) dx \) and \( c = \int_{NMEB} J \cdot ds \). Term \( c \) depends on \( u(NM), u(ME) \) and \( u(EB) \). We fix these utilities and study \( u(BN) \).

When we change \( u(BN) \), the position of \( ME \) will change, but \( u(ME) \) is always zero because \( ME \) is defined to be the boundary of the \((0, 0)\) region.

We manipulate the utility of \( BN \) under the convex constraint. As a result, we manipulate the height and the gradients of the plane.

Note that, no matter how we manipulate, the buyer’s utility function on the full domain is still the upper envelope of a set of linear utility planes. The convex constraint is preserved through any such manipulation. The IC constraint is equivalent to the constraint that the utility function is convex. So the IC constraint is preserved through any such manipulation.

Let \( R_u(u(x_B, y), y) \) denote the partial derivative with respect to \( u(x_B, y) \):

\[
f_2(y)[x_B f_1(x_B) - \int_{x_B - u(x_B, y)}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y))) dx]
\]

Let \( v(l) = x_B f_1(x_B) - \int_{x_B - u(x_B, y)}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y))) dx \). Under Condition 3, \( 3 + PR(f_1(x)) + PR(f_2(y)) \geq 0 \) is a constant, \( v \) is an increasing function of \( l \). When \( x_B f_1(x_B) - v(u(x_B, y)) > 0 \), \( R_{EBN_M} \) decreases as \( u(x_B, y) \) increases. In Fig.5, when \((x_B, y)\) moves from \( B \) to \( N \), \( u(x_B, y) \) weakly increases, \( v(x_B - u(x_B, y)) \) weakly decreases. There are 3 cases:

1. \( v(x_B - u(x_B, y)) > 0, y \in [y_B, y_N] \),
2. \( v(x_B - u(x_B, y)) < 0, y \in [y_B, y_N] \),
3. \( \exists y' \in [y_B, y_N], v(x_B - u(x_B, y')) = 0, v(x_B - u(x_B, y)) \geq 0, y \in [y_B, y'] \) and \( v(x_B - u(x_B, y)) \leq 0, y \in [y', y_N] \).

(1) can be regarded as a special case of (3) by setting \( y' = y_B \), because \( u(b, y), y \in [y_B, y_N] \) must be also as large as possible. (2) can be regarded as an special case of (3) by setting \( y' = y_N \), because \( u(x_B, y), y \in [y_B, y_N] \) must be as small as possible.
So it is without loss of generality to restrict attention to case (3). Since $K$ is randomly chosen, set $y' = y_K$. The revenue increases as $u(KN)$ decreases and revenue increases as $u(KB)$ increases.

Consider what the optimal function $u(BN)$ is, given $u$ is convex and fixed values of $u(B)$, $u(K)$ and $u(N)$. The optimal $u(BN)$ comprises of two lines: the straight line (with extended line) across points $(y_B, u(B)) (y_K, u(K))$, and the straight line across point $(y_N, u(N))$ with slope $q_2(N)$.

Since $u$ is convex,
\[ u'(x_B, y) \leq u'(x_B, y_N) = q_2(N), y \in [y_K, y_N] \]
\[ u'(x_B, y) \geq u'(x_B, y_K) \geq \frac{u(K) - u(B)}{y_K - y_B}, y \in [y_K, y_N] \]

To maximize $u(BK)$, $u(x_B, y)$ must have the same derivative $\frac{u(K) - u(B)}{y_K - y_B}$. Now consider $u(KN)$, $u(x_B, y) \geq u(K) + (y - y_K) * \frac{u(K) - u(B)}{y_K - y_B}$, and $u(x_B, y) = u(N) - \int_{y_K}^{y_N} q_2(x_B, y)dy \geq u(N) - q_2(N) * (y_N - y)$. By our construction, $u(KN)$ achieves the lower bound $u(x_B, y) = \max\{u(K) + (y - y_K) * \frac{u(K) - u(B)}{y_K - y_B}, u(N) - q_2(N) * (y_N - y)\}$. □

Lemma 5.4. There exists an optimal utility function such that $u(ND)$ is piecewise linear with at most 2 pieces, when there is no point on $CA$ choosing menu item $(1, *)$.

Proof. The revenue obtained in region $MNDG$ is
\[ R_{MNDG} = \int_{MNDG} J \cdot \hat{\mathbf{n}} ds - \int_{MNDG} \Delta(z) u(z) dz \]
\[ = \int_{DGMN} J \cdot \hat{\mathbf{n}} ds + \int_{ND} J \cdot \hat{\mathbf{n}} ds - \int_{y_N}^{y_D} \int_{x_M}^{x_N} \Delta(x, y) u(x, y) dx dy \]
\[ = c + \int_{y_N}^{y_D} [x_B u(x_B, y) f_1(x_B) f_2(y) - \int_{x_M}^{x_N} \Delta(x, y) u(x, y) dx] dy \]
\[ = c + \int_{y_N}^{y_D} R(u(x_B, y), y) dy \]

\[4\text{Assuming differentiability of } u \text{ is ok, since } u \text{ is differentiable almost everywhere.} \]
\[ c = \int_{DGMN} J \cdot \hat{n} ds \] depends on \( u(DG), u(GM) \), and \( u(MN) \). We fix these utilities and study \( u(DN) \). Let \( R(u(x_B, y), y) = x_B u(x_B, y) f_1(x_B) f_2(y) - \int_{x_M}^{x_N} \triangle(x, y) u(x, y) dx \). Pick a random point \( P \) on \( ND \), draw a horizontal line across \( P \) that intersects curve \( MLQ \) at \( J \), intersects segment \( MG \) at \( R \). In region \( CSMLID \) point \((x, y)\) gets item 2 deterministically and we assume it gets item 1 with probability \( q_1(x) \).

\[
R(u(x_B, y_J), y_J) = u(x_B, y_J) x_B f_1(x_B) f_2(y_J) - \int_{x_M}^{x_J} u(x, y_J) \triangle(x, y_J) dx - \int_{x_J}^{x_B} u(x, y_J) \triangle(x, y_J) dx
\]

\[
= [u(R) + \int_{x_M}^{x_J} q_1(x) dx + x_B - x_J] x_B f_1(x_B) f_2(y_J) - \int_{x_M}^{x_J} [u(R) + \int_{x_B}^{x} q_1(l) dl] \triangle(x, y_J) dx
\]

\[
- \int_{x_J}^{x_B} [u(R) + \int_{x_M}^{x_J} q_1(l) dl + x - x_J] \triangle(x, y_J) dx
\]

We want to see the relationship between \( u(x_B, y_J) \) and \( R(u(x_B, y_J), y_J) \) when \( y_J \) is fixed. We use the intermediate variable \( x_J \). Given \( y_J \) and \( u(CD) \) fixed, we study the boundary between the \((*, 1)\) region and the \((1, *)\) region on horizontal line \( y = y_J \). Note that \( q_1 < 1 \) in the \( GMJLQ \) region, and \( q_1 = 1 \) in the \( MJLQDN \) region. The idea is when \( u(x_B, y_J) \) increases, the corresponding plane goes up, the intersection point \((x_J, y_J)\) goes left, i.e. \( x_J \) decreases. The formula that reflects this relation is

\[
u(x_B, y_J) = u(R) + \int_{x_M}^{x_J} q_1(x) dx + x_B - x_J = u(R) + x_B - \int_{x_M}^{x_J} (1 - q_1(x)) dx
\]

We have \( \frac{\partial u}{\partial x_J}(x_B, y_J) \leq 0 \).

\[
\frac{\partial R}{\partial x_J}(u(x_B, y_J), y_J)
\]

\[
= (q_1(x_J) - 1) x_B f_1(x_B) f_2(y_J) - \int_{x_J}^{x_B} (q_1(x_J) - 1) \triangle(x, y_J) dx
\]

\[
= (q_1(x_J) - 1) f_2(y_J)[x_B f_1(x_B) - \int_{x_J}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y_J))) dx]
\]

\[
= (q_1(x_J) - 1) f_2(y_J)v(x_J)
\]
In the last equality, we reset \( v(x_J) = x_B f_1(x_B) - \int_{x_J}^{x_B} f_1(x) (3 + PR(f_1(x)) + PR(f_2(y_J))) \, dx \). Since \( q_1(x_J) - 1 \leq 0 \), \( sgn(v(x_J)) = -sgn(\frac{\partial R}{\partial x_J}) \). Thus \( sgn(v(x_J)) = -sgn(\frac{\partial R}{\partial x_J}) = -sgn(\frac{\partial R}{\partial u} \frac{\partial u}{\partial x_J}) = sgn(\frac{\partial R}{\partial u}) \).

The second equality is because \( y_J \) is fixed. Under Condition 2, \( 3 + PR(f_1(x)) + PR(f_2(y)) \geq 0 \), \( v(x_J) \) is weakly monotone increasing in \( x_J \). By (3) and (4) of Lemma 5.1, while \( P \) moves up, the intersection \( J \) moves towards the right, i.e., \( x_J \) weakly increases. Then \( v(x_J) \) weakly increases. So if \( v(x_J) \) switches sign, it must switch sign from minus to plus. So \( R_u(u(x_B, y_J), y_J) \) can only switch sign from minus to plus.

There are three cases for the sign of \( R_u(u(x_B, y_J), y_J) \). It is similar to the proof in Lemma 5.3. WLOG, we can assume that \( v(x_J) = 0 \). Thus \( v(x) \geq 0, x \in [x_J, x_D] \) and \( v(x) \leq 0, x \in [x_A, x_J] \). Therefore \( R_u(u(x_B, y_J), y_J) \geq 0, y \in [y_P, y_D], R_u(u(x_B, y_J), y_J) \leq 0, y \in [y_N, y_P] \). So \( R_{MNDC} \) increases as \( u(NP) \) decreases and increases as \( u(PD) \) increases. With fixed \( u(N) \), \( u(P) \), and \( u(D) \), by convexity, optimal \( u(ND) \) comprises of two lines: the straight line (with extended line) across points \((y_P, u(P))(y_D, u(D))\), and straight line across point \((y_N, u(N))\) with slope \( q_2(N) \). □

With these two lemmas, we are able to prove Theorem 5.2.

Proof.

![Figure 6: Utility function on BD.](image-url)
Case 1: there is no point on CA choosing menu item \((1, *)\).

We sum up the conclusions drawn on different segments of \(BD\) and settle the final shape of \(u(BD)\), subject to the convex constraint and the facts that \(u(B), u(K), u(P)\) and \(u(D)\) are fixed. In Fig. 7, the black solid line is an arbitrary convex utility function. By Lemma 5.3 and 5.4, the optimal \(u(BD)\) consists of at most 3 pieces: the straight line (with extended line) across points \((y_B, u(B))(y_K, u(K))\), the straight line (with extended line) across points \((y_P, u(P))(y_D, u(D))\), and the straight line across point \((y_N, u(N))\) with slope \(q_2(N)\). In fact, we can prove an even stronger result: it turns out that the line across point \((y_N, u(N))\) is unnecessary. The remainder of the proof is to confirm this claim.

In Fig. 7, the red dashed utility consists of two parts: the straight line across points \((y_B, u(B))\) and \((y_K, u(K))\), and the straight line across points \((y_P, u(P))\) and \((y_D, u(D))\). We denote the “original” utility function to be the black arbitrary convex utility function. We denote the “new” utility function to be the red dashed utility function. We prove that the revenue based on the new utility function is greater than or equal to the revenue based on the black original utility. Therefore when \(u(B), u(K), u(P), u(D)\) and their coordinates on the \(y\)-dimension \(y_B, y_K, y_P, y_D\) are fixed, the optimal utility function must be of the shape portrayed as the red dashed line.

First, we study what the graph representation looks like. Since \(u(Q), u(J), u(M)\) and \(u(W)\) remain the same, they are still on the boundaries (see Fig. 5):

- \(M\) is on the boundary between the \((*, 1)\) region and the \((0, 0)\) region.
- \(W\) is on the boundary between the \((1, *)\) region and the \((0, 0)\) region.
- \(J\) is on the boundary between the \((1, *)\) region and the \((*, 1)\) region.
- \(Q\) is on the boundary between the \((*, 1)\) region and the \((1, 1)\) region.

Let \(M'\) denote the new intersection of the three parts: \((0, 0), (*, 1), \) and \((1, *)\). The new boundary between \((1, *)\) and \((*, 1)\) is the dashed line \(QJM'\). The new boundary between \((*, 1)\) and \((0, 0)\) is the dashed line \(SMM'\). The new boundary
between \((1,\ast)\) and \((0,0)\) is the dashed line \(M'WE\). Compared to the original utility, the new utility \(u(PD)\) and \(u(BK)\) weakly increases, and \(u(KP)\) weakly decreases, \(M'\) must lie in the \(MWKN\) region, dashed \(QJ\) lies in the \(GMLQ\) region, dashed \(WE\) lies in the \(SMWE\) region. Draw horizontal line through \(M'\) which intersects \(BD\) at \(N'\), and intersects the extended line of \(GM\) at \(O\).

With fixed \(u(CD)\), given the boundary of \((1,\ast)\), we can calculate \(u(BD)\) as follows: say \((x,y)\) is on the boundary between \((1,\ast)\) and \((\ast,1)\), then \(u(x_B,y) = u(x,y) + x_B - x = u(x,y_C) + y - y_C + x_B - x\). So, we can define \(u(BD)\) according to their boundary. Let the original utility and revenue function based on the boundary \(QLJMWE\) be \(u_1\) and \(R_1\). Let the new utility and revenue function based on the boundary \(QJM'W\) and dashed \(WE\) be \(u_3\) and \(R_3\). Our goal is to prove \(R_1 \leq R_3\).

\[
R_{DGMWEB}^1 \leq R_{DGMWEB}^3
\]

Since \(R_{DGMWEB}^1 = R_{DGMWK}^1 + R_{WEBK}^1\) (here \(WE\) is the solid line) and \(R_{DGMWEB}^3 = R_{DGM'MWK}^3 + R_{WEBK}^3\) (here, \(WE\) is the dashed line), it suffices to prove \(R_{DGMWK}^1 \leq R_{DGM'MWK}^3\) and \(R_{WEBK}^1 \leq R_{WEBK}^3\) separately.

\[
R_{WEBK}^1 = \int_{WK} J^1 \cdot \hat{n} ds + \int_{EB} J^1 \cdot \hat{n} ds + \int_{y_B}^{y_K} u_1(x_B,y)x_B f_1(x_B)f_2(y)dy - \int_{y_B}^{y_K} \int_{x_B-u(x_B,y)}^{x_B} u_1(x,y)\Delta(x,y)dxdy
\]

\[
= \int_{WK} J^3 \cdot \hat{n} ds + \int_{EB} J^3 \cdot \hat{n} ds + \int_{y_B}^{y_K} R(u_3(x_B,y),y)dy
\]

\[
\leq \int_{WK} J^3 \cdot \hat{n} ds + \int_{EB} J^3 \cdot \hat{n} ds + \int_{y_B}^{y_K} R(u_3(x_B,y),y)dy
\]

\[
= R_{WEBK}^3
\]

Here \(J^i = zw^i f(z), i = 1, 3\) and we use the same \(R(u(x_B,y),y)\) as that in Lemma 5.3. What still remains to show is \(R_{DGMWK}^1 \leq R_{DGM'MWK}^3\).

In order to prove this claim, we introduce an intermediate utility and revenue as follows. Let the intermediate utility and revenue function based on the boundary \(QLJM'MW\) and dashed \(WE\) be \(u_2\) and \(R^2\). In the intermediate case,
\( u(N'N) \) is trimmed so that \( MM' \) is the boundary between regions \((1, *) \) and \((0, 0) \). For any point \((x, y)\) on \( MM' \), we can calculate the utility of corresponding point \((x_B, y)\) as \( u(x_B, y) = u(x, y) + x_B - x = u(x, y_C) + y - y_C + x_B - x \). In fact, according to Lemma 5.1, this case cannot happen and \( u^2(BD) \) does not retain convexity any more. It is important to note that, here, we are only concerned with \( R^2 \), not the feasibility of \( u^2 \). That is, we use \( R^2 \) as a benchmark to facilitate the comparison between \( R^1 \) and \( R^3 \).

Now we show \( R^1_{DGMWK} \leq R^3_{DGM'M'WK} \):

\[
R^1_{DGMWK} = R^1_{DGMN} + R^1_{MWKN} \leq R^1_{DGMN} + R^2_{MM'WKN} = R^2_{DGMN} + R^3_{MM'N'N} + R^2_{MM'WKN'} = R^2_{DGM'M'N'} + R^2_{MM'WKN'} \leq R^3_{DGM'M'N'} + R^3_{MM'WKN'} = R^3_{DGM'M'WKN'}
\]

The first inequality follows from a similar proof of that in \( R^1_{WEBK} \leq R^3_{WEBK} \). We have \( u^2(KN) \leq u^1(KN) \). Using the same \( R(u(x_B, y), y) \) as that in Lemma 5.3, we get \( \int_{y_k}^{y_N} R(u^1(x_B, y), y)dy \leq \int_{y_k}^{y_N} R(u^2(x_B, y), y)dy \). Then we derive \( R^1_{MWKN} \leq R^2_{MM'WKN} \).

In order to show the second inequality, introduce \( u^{22} \) and \( u^{33} \) as follows:

\[
u^{22}(x, y) = \begin{cases} 
  u^2(x, y) & (x, y) \in DGM'M'N' \\
  u^2(x, y_C) + y - y_C & (x, y) \in MM'O
\end{cases}
\]

\[
u^{33}(x, y) = \begin{cases} 
  u^3(x, y) & (x, y) \in DGM'M'N' \\
  u^3(x, y_C) + y - y_C & (x, y) \in MM'O
\end{cases}
\]

Then we have

\[
R^2_{DGM'M'} - R^3_{DGM'M'}
\]
\begin{align*}
&= \int_{DGMM'} \mathcal{J}^2 \cdot \hat{n} ds + \int_{N'} \mathcal{J}^2 \cdot \hat{n} ds - \int_{y_{MM'}} \int_{x_{MM'}} \triangle(x, y) u^{22}(x, y) dxdy \\
&\quad + \int_{z_{MM'O}} \triangle(z) u^{22}(z) dz - R_{DGMM'}^3 \\
&= \int_{y_{N'}} [x_B u^{22}(x_B, y) f_1(x_B) f_2(y)] - \int_{x_{M}} \triangle(x, y) u^{22}(x, y) dx]dy \\
&\quad - \int_{y_{N'}} [x_B u^{33}(x_B, y) f_1(x_B) f_2(y)] - \int_{x_{M}} \triangle(x, y) u^{33}(x, y) dx]dy \\
&= \int_{y_{N'}} [R(u^{22}(x_B, y), y) - R(u^{33}(x_B, y), y)]dy 
\end{align*}

In the last equality, we use the same \(R(u(x_B, y), y)\) as in Lemma 5.4. In the following, we show that the right-hand side of Equation (2) is non-positive. According to Lemma 5.4,

\begin{align*}
R_u(u(x_B, y), y) &\geq 0 \quad y \in [y_J, y_Q], u(x_B, y) \geq u(J) \\
R_u(u(x_B, y), y) &\geq 0 \quad y \in [y_{MM'}, y_J], u(x_B, y) \leq u(J)
\end{align*}

Because \(u^{22}(x_B, y) \leq u^{33}(x_B, y) \quad y \in [y_P, y_D]\) and \(u^{22}(x_B, y) \geq u^{33}(x_B, y) \quad y \in [y_{MM'}, y_P]\), we have \(R(u^{22}(x_B, y), y) \leq R(u^{33}(x_B, y), y), \quad y \in [y_{MM'}, y_P]\). So Equation (2) \(\leq 0\), then \(R_{DGMM'}^2 \leq R_{DGMM'}^3\).

Up to now, we have proved that red dashed utility function on \(BD\) segment indeed yield the highest revenue subject to fixed \(u(B), u(K), u(P), \) and \(u(D)\) (Fig. 7). In other words, \(u(BD)\) is piecewise linear with two pieces.

Case 2: when there is some point on \(CA\) choosing menu item \((1, \ast)\).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig7.png}
\caption{The optimal allocation when there is no point on curve BME chooses allocation menu item \((1,1)\).}
\end{figure}
denotes the \((*, 1)\) region and \(ASWE\) denotes the \((0, 0)\) region. \(MS\) denotes the interval on \(AC\) that choose allocation \((1, *)\). \(MN\) and \(SI\) are horizontal lines.

Consider the region \(MSIN\).

\[
R_{SMNI} = \int_{SM} J \cdot \hat{\nu}ds + \int_{MN} J \cdot \hat{\nu}ds + \int_{N1} J \cdot \hat{\nu}ds + \int_{IS} J \cdot \hat{\nu}ds - \int_{SMNI} \triangle(z)u(z)dz
\]

\[
= c + \int_{ys}^{ym} u(x_B, y)x_Bf_1(x_B)f_2(y)dy - \int_{ys}^{ym} (u(x_B, y) + x_A - x_B)x_Af_1(x_A)f_2(y)dy - \int_{ys}^{ym} \int_{x_A}^{x_B} \triangle(x, y)(u(x_B, y) + x - x_B)dxdy
\]

\[
= c + \int_{ys}^{ym} R(u(x_B, y), y)dy
\]

Here, \(c = \int_{S1} J \cdot \hat{\nu}ds + \int_{MN} J \cdot \hat{\nu}ds\). When we fix \(u(N)\) and \(u(I)\), \(c\) is constant. Let \(R = u(x_B, y)x_Bf_1(x_B)f_2(y) - (u(x_B, y) + x_A - x_B)x_Af_1(x_A)f_2(y) - \int_{x_A}^{x_B} \triangle(x, y)(u(x_B, y) + x - x_B)dxdy\) and \(R_u(u(x_B, y), y)\) denote the partial derivative w.r.t. \(u(x_B, y)\).

\[
R_u(u(x_B, y), y) = [x_Bf_1(x_B) - x_Af_1(x_A) - \int_{x_A}^{x_B} f_1(x)[3 + PR(f_1(x)) + PR(f_2(y))]dx]f_2(y)
\]

The sign of it is unchanged. When it is plus, \(R_{SINM}\) increases as the utility on segment \(IN\) decreases and vice versa.

In region \(CMND\), we can use the same argument as in Lemma 5.4. When there exists a point \(P\) such that \(R_{MNDC}\) increases as \(u(NP)\) decreases and increases as \(u(PD)\) increases, we have \(v(x_j) = 0\) and

\[
v(x_j) = x_Bf_1(x_B) - \int_{x_j}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y)))dx
\]

\[
\geq x_Bf_1(x_B) - \int_{x_A}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y)))dx
\]

\[
> R_u(u(x_B, y), y)
\]

In region \(ABIS\), we can use the same argument as Lemma 5.3. When there exists a point \(K\) such that \(R_{SEBI}\) is increasing as \(u(KI)\) decrease and is
increasing as $u(BK)$ increase, we have $v(x_B - u(K)) = 0$ and

$$v(x_B - u(K)) = x_B f_1(x_B) - \int_{x_B - u(K)}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y)))dx$$

$$= x_B f_1(x_B) - \int_{x_B - u(K)}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y)))dx$$

i.e. $v(x_B) = x_B f_1(x_B) - \int_{x_B - u(K)}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y)))dx$

$$\geq x_B f_1(x_B) - \int_{x_A}^{x_B} f_1(x)(3 + PR(f_1(x)) + PR(f_2(y)))dx$$

$$> R_u(u(x_B), y)$$

When $R_u(u(x_B), y) \leq 0$, wlog, we can assume that points $K$ and $P$ exist.

5 with fixed $u(B), u(K), u(P)$ and $u(D)$, $u(DP)$ and $u(BK)$ should be as large as possible, $u(KP)$ should be as small as possible. We construct the new utility function same as the dashed utility in Case 1 (shown in Fig.7). $u(BD)$ is piecewise linear with two pieces.

When $R_u(u(x_B), y) > 0$, points $K$ and $P$ do not exist. The total revenue increases as $u(ND)$ increases, $u(IN)$ increases and $u(BI)$ increases. This means that in all three intervals: where the $(0, 0)$ region is on the left of the $(1, *)$ region, where the $(1, *)$ region is on the left of the $(1, *)$ region, and where edge $AC$ is on the left of the $(1, *)$ region, $u(BD)$ should be as large as possible with fixed $u(B)$ and $u(D)$. We let $u(BD)$ be a straight line and the total revenue increases as a result.

So no matter what $R_u(u(x_B), y)$ is, $u(BD)$ is piecewise linear with two pieces at most.

Combining the two cases, no matter whether there is a point on $CA$ choosing menu item $(1, *)$, $u(BD)$ is piecewise linear with two pieces at most. Same for $u(CD)$.

We have showed that region $CSEBD$ consists of 4 menu items. In other words, the whole mechanism consists of 5 menu items. Say, points on $BD$ segments choose menu items $(1, \alpha, t_\alpha), (1, \gamma, t_\gamma)$, $\alpha \leq \gamma$. Points on $CD$ choose

\footnotesize
5For example, if $K$ does not exist, we take $K$ as $B$ or $I$.

31
menu items $(\beta, 1, t_\beta), (\theta, 1, t_\theta)$, $\beta \leq \theta$. What remains to show is that the top-right two regions both allocate with probabilities $(1, 1)$ thus are in fact a unique region.

Manelli and Vincent (2007, Theorem 16) state that in the optimal mechanism, there must exist a segment that chooses allocation $(1, 1)$. By Lemma 5.1, the $(1, 1)$ region is on the top-right corner of the rectangle. So, in the optimal mechanism, there are slopes of utility lines equal to 1 in both $BD$ and $CD$. Hence, $\gamma = \theta = 1$. $BD$ and $CD$ share the same menu item $(1, 1)$.

\[
\begin{array}{c|c|c}
q_1 & q_2 & t \\
0 & 0 & 0 \\
1 & \alpha & t_\alpha \\
\beta & 1 & t_\beta \\
1 & 1 & t_1 \\
\end{array}
\]

Hart and Nisan (2012, Theorem 1 and Lemma 14) state that bundling 4-approximates the optimal revenue for general two-item setting. As an application of Theorem 5.2, we obtain a better lower bound for the bundling auction.

**Corollary 5.5.** Under Conditions 2 and 3, bundling 3-approximates the optimal revenue.

**Proof.** The revenue of an optimal mechanism with 3 non-zero menu items is less than or equal to the sum of revenues of 3 mechanisms, each of which has only 1 non-zero menu items. Since bundling is optimal among all mechanisms that contain only 1 non-zero menu item, it is not worse than any of these three mechanisms. Consequently bundling gives a 3-approximation to the optimal revenue.\(^6\)

\[^6\text{Compare any mechanism} \,(0, 0, 0), (q_1, q_2, t) \text{ and the bundling mechanism} \,(0, 0, 0), (1, 1, t). \text{ First of all, notice that two non-zero menu items have the same payment. It should also be} \]

32
Two cases where the optimal mechanism contains \( \leq 3 \) items

In fact, Conditions 1, 2 and 3 have intersections. When all conditions are satisfied, the revenue does not depend on the utilities on inner points any more. In this case, we obtain a condition under which there is an optimal mechanism that contains at most 3 menu items.

**Corollary 5.6.** For \( f_1(x) = s_1 x^{i_1}, f_2(y) = s_2 y^{i_2}, s_1, s_2 > 0, i_1 + i_2 = -3 \), there is an optimal mechanism that contains at most 3 menu items, thus bundling gives a 2-approximation of the optimal revenue.

**Proof.** By Formula (1), we can see that the revenue only depends on the utility on the boundaries. According to Theorem 5.2, there are at most 4 menu items and both \( u(BD) \) and \( u(CD) \) are piecewise linear with two pieces. Suppose otherwise that there are 4 different menu items. Then one could raise up the plane of the \((1,1)\)-item uniformly, i.e. expand the top-right region, until one of \( u(BD), u(CD) \) becomes a straight line, i.e. the top-right region covers part of \( AB \) or \( AC \). Note that this procedure will not change \( u(AB) \) or \( u(AC) \) since it will terminate as long as the \((1,1)\)-item reaches \( AB \) and \( AC \). Moreover, this procedure increases \( u(BD) \) and \( u(CD) \) while maintaining convexity. So the new utility function corresponds to a strictly higher revenue, which contradicts the fact that \( u \) is optimal. \( \square \)

Following a proof similar to Theorem 5.2, we obtain another condition under which 3 menu items are enough. Note that this condition does not impose constant power rate, thus is not a special case of Condition 3.

**Condition 4:** \(-2 \leq PR(f_1(x)) \leq y_A f_2(y_A) - 2, \forall x \) and \(-2 \leq PR(f_2(y)) \leq x_A f_1(x_A) - 2, \forall y.\)

clear that for any type that is willing to choose \((q_1, q_2, t)\) in the first mechanism must also be willing to choose \((1, 1, t)\) in the bundling mechanism. As a result, the items are sold more often in the bundling mechanism than in the first mechanism, with the same payment if sold. In other words, bundling is optimal among all such mechanisms.
Theorem 5.7. Under Conditions 2 and 4, there is an optimal mechanism which contains at most 3 menu items, thus bundling gives a 2-approximation of the optimal revenue.

Proof.

![Figure 8: Optimal mechanism with 3 menu items.](image)

We consider two cases:

Case 1: There is no type on edge $AC$ that chooses the first item with probability 1 and chooses the second item randomly, i.e., $(1,*)$ region does not touch edge $AC$.

Look at Fig. 8, the allocations in region $SMJDC$ are in the form $(*,1)$. When $y \in [y_B, y_N]$, we have $u(x_B, y) \leq x_B - x_A$, Lemma 5.3 implies $R_u(u(x_B, y), y) \geq f_2(y)\left[x_B f_1(x_B) - \int_{x_A}^{x_B} f_1(x)[3 + PR(f_1(x)) + PR(f_2(y))]dx\right]$.

When $y \in [y_N, y_D]$, we have $x_J \geq x_A$, Lemma 5.4 implies $v(x_J) \geq x_B f_1(x_B) - \int_{x_A}^{x_B} f_1(x)[3 + PR(f_1(x)) + PR(f_2(y))]dx$.

We have $x_B f_1(x_B) - \int_{x_A}^{x_B} f_1(x)[3 + PR(f_1(x)) + PR(f_2(y))]dx$.

---

7 In Lemma 5.3, we use Condition 3 after the definition of $R_u(u(x_B, y), y)$.
8 In Lemma 5.4, we have not used Condition 3 at all.
After this operation, points \( u \) as above, the total revenue will weakly increase. Now there are 2 non-zero methods as is shown in Theorem 5.2, the total revenue will weakly increase. Theorem 5.2, we can prove that with fixed \( u(B) \), \( u(N) \), and \( u(D) \), the optimal utility \( u(ND) \) and \( u(BN) \) are straight lines. Using similar methods as shown in Theorem 5.2, we can prove that with fixed \( u(B) \) and \( u(D) \), the optimal utility function is \( u(x_B, y) = u(B) + q_{23}(y - y_B) \) \( y \in [y_B, y_D] \), where \( q_{23} = \frac{u(D) - u(B)}{y_D - y_B} \). But we should notice that this operation may change \( u(AC) \). This leads to two further cases: (1) the operation changes \( u(AC) \); (2) otherwise. It can be seen that case (1) is strictly more general, so we only need to consider case (1).

Draw a vertical line through \( S \) across \( BD \) at \( P \) and fix \( u(B) \) and \( u(D) \). Instead of letting \( u(BD) \) be a straight line, we let \( u(BD) \) be piece-wise linear with one break point \( u(P) \) in order not to change \( u(AC) \). Let \( u(P) = x_B - x_A \),

\[
u(x_B, y) = u(P) + \frac{y - y_P}{y_D - y_P} \cdot (u(D) - u(P)) \quad y \in [y_P, y_D]
\]

\[
u(x_B, y) = u(B) + \frac{y - y_B}{y_P - y_B} \cdot (u(P) - u(B)) \quad y \in [y_B, y_P]
\]

After this operation, points \( u(S), u(P), u(D) \) are in a new plane. So the boundary of the \((*, 1)\) region, curve \( MJD \), moves into the \( CSD \) region. Using similar methods as is shown in Theorem 5.2, the total revenue will weakly increase.

Then we set

\[
u(x, y_C) = u(C) + \frac{x - x_C}{x_D - x_C} \cdot (u(D) - u(C)) \quad x \in [x_C, x_D]
\]

After this operation, points \( u(C), u(S), u(D) \) are in a new plane. The boundary of the \((*, 1)\) region, originally curve \( MJD \), becomes the straight line \( SD \). So this operation will not affect the utilities on segment \( AB \). Using the same argument as above, the total revenue will weakly increase. Now there are 2 non-zero
menu items chosen in region CSPD and 1 non-zero menu item chosen in region SMEBP. Keeping $u(S)$ fixed, $x_S + y_S - u(S)$ is the largest payment from point $S$. Note that all three menu items are all the best choices for point $S$. The payment in all menu items is bounded by $x_S + y_S - u(S)$.

We add a new menu $(1, 1, x_S + y_S - u(S))$. The corresponding plane goes through point $S$, $C$, $P$. It is higher than the current utility surface in the CSPD region, and lower in the ASBP region. The types in the CSPD region will choose this new menu. The types in the ASBP region will not choose this new menu. Then the types in the CSPD region have weakly higher payment, and the total revenue weakly increases. Thus there are two non-zero menu items in total: $(1, 1, t)$ chosen by the CSPD region and $(1, \alpha, t_1)$ chosen by the SMEBP region.

Case 2: There is some type on edge $AC$ that chooses the first item with probability 1 and choose the second item randomly, i.e., the $(1, *)$ region touches edge $AC$. Let $M$ denote the boundary between the $(*)$, 1) region and the $(1, *)$ region.

Using similar arguments as Case 1, we can prove the following. With $u(M)$ fixed, we can weakly increase the revenue by adding a new menu $(1, 1, x_M + y_M - u(M))$. With $u(B)$ and $u(P)$ fixed, we can weakly increase the revenue by setting $u(PB)$ as a straight line. Next, we consider the SMNP region. Now Condition “$PR(f_1(x)) \leq y_A f_2(y_A) - 2, \forall x$” is not enough, because when

Figure 9: Optimal mechanism with 3 menu items.
we do the same manipulation for \( u(BD) \), \( u(CS) \) will change at the same time. According to this, we have the following equation:

\[
R_{SPNM} = \int_{SP} J \cdot \hat{n} ds + \int_{MN} J \cdot \hat{n} ds + \int_{PN} J \cdot \hat{n} ds \\
+ \int_{SM} J \cdot \hat{n} ds - \int_{SPNM} \triangle(z) u(z) dz \\
= c + \int_{YS}^{YM} u(x_B, y) x_B f_1(x_B) f_2(y) dy \\
- \int_{YS}^{YM} (u(x_B, y) + x_A - x_B) x_A f_1(x_A) f_2(y) dy \\
- \int_{YS}^{YM} \int_{x_A}^{x_B} \triangle(x, y) (u(x_B, y) + x - x_B) dx dy \\
= c + \int_{YS}^{YM} R(u(x_B, y), y) dy
\]

Here, \( c = \int_{SP} J \cdot \hat{n} ds + \int_{MN} J \cdot \hat{n} ds \). When we fix \( u(N) \) and \( u(P) \), \( c \) is constant. Let \( R = u(x_B, y) x_B f_1(x_B) f_2(y) - (u(x_B, y) + x_A - x_B) x_A f_1(x_A) f_2(y) - \int_{x_A}^{x_B} \triangle(x, y) (u(x_B, y) + x - x_B) dx \) and \( R_u(u(x_B, y), y) \) denote the partial derivative wrt. \( u(x_B, y) \).

\[
R_u(u(x_B, y), y) \\
= [x_B f_1(x_B) - x_A f_1(x_A) - \int_{x_A}^{x_B} f_1(x) |3 + PR(f_1(x)) + PR(f_2(y))| dx] \cdot f_2(y) \\
= [-2 - PR(f_2(y))] \cdot f_2(y) \leq 0
\]

This means that \( R_{SPNM} \) increases as utilities on segment \( PN \) decreases and decreases as utilities on segment \( PN \) increases.

As shown in Fig. 10, the new mechanism with the red dashed line \( u(BD) \) yields higher revenue than the original mechanism with the black solid line \( u(BD) \). So there are two non-zero menu items in total: \( (1, 1, t) \) and \( (1, \alpha, t_1) \). □

### 5.3. Optimal mechanisms for i.i.d. monotone increasing power rate

The requirement of the power rate to be constant might be restrictive. If one relaxes this requirement to monotone increasing power rate, one only needs
Condition 5: $PR(f_i(x)), i = 1, 2$, is weakly monotone increasing.

Theorem 5.8. Under Conditions 2 and 5, if $f_1 = f_2$, there is an optimal mechanism that consists of at most 6 menu items.

The general form of the optimal mechanism is shown in Fig. 11. It is without loss of generality to restrict attention to symmetric mechanisms (Maskin and Riley, 1984, section 1). Let $AD$ intersect $SE$ at point $M$. In region $ASME$, the seller keeps both items. Item 2 is sold deterministically in $CSMD$ and item...
1 is sold deterministically in \( MEBD \). Let the allocation rule on point \((x, y)\) in \( CSMD \) be \((q_1(x), 1)\). Similar to the proof of Theorem 5.2, we start with the following lemma.

**Lemma 5.9.** There is an optimal utility function such that \( u(ND) \) is piecewise linear with at most 2 pieces.

**Proof.** Since point \( M \) is on the \((0, 0)\) part, \( u(M) = 0 \). For point \((x, y)\) in region \( NMD \), \( u(x, y) = u(M) + y - y_M + \int_{x_M}^{x} q_1(l)dl = y - y_M + \int_{x_M}^{x} q_1(l)dl \).

Rewrite the revenue formula for region \( NMGD \) as follows,

\[
R_{NMGD} = \oint_{NMGD} \mathcal{J} \cdot \hat{n}ds - \int_{NMGD} \triangle(z)u(z)dz
\]

\[
= 2 \int_{x_N}^{x_D} y_C u(x, y_C) f_1(y_C) f_1(x)dx - 2 \int_{N_M} \mathcal{J} \cdot \hat{n}ds
\]

\[
-2 \int_{x_N}^{x_D} \int_x^{y_C} u(x, y) \triangle(x, y)dydx
\]

\[
= 2 \int_{x_N}^{x_D} x_B[y_C - y_M + \int_{x_N}^{x} q_1(l)dl] f_1(x_B) f_1(x)dx
\]

\[
-2 \int_{x_N}^{x_D} \int_x^{y_C} [y - y_M + \int_{x_N}^{x} q_1(l)dl] \triangle(x, y)dydx - 2 \int_{N_M} \mathcal{J} \cdot \hat{n}ds
\]

\[
= 2 \int_{x_N}^{x_D} q_1(x)(\int_{x}^{x_D} [x_B f_1(x_B) f_1(l) - \int_{l}^{x_D} \triangle(l, y)dy]dl)dx + C_1 \quad (3)
\]

\[
= 2 \int_{x_N}^{x_D} q_1(x)v(x)dx + C_1
\]

In Equation (3), we put the terms that are related to \( q_1 \) together, and the other terms together denoted by \( C_1 \) which only depends on \( y_M \). When \( y_M \) is fixed, \( C_1 \) is a constant. In the last equality, let \( v(x) = \int_{x}^{x_D} [x_B f_1(x_B) f_1(l) - \int_{l}^{x_D} \triangle(l, y)dy]dl \), thus it is independent of \( q_1 \).

To find out the optimal \( q_1 \), divide \([x_N, x_D]\) into several intervals according to \( v(x) \). Since \( v(x) \) is a continuous function, we can assume in intervals \( v(x) \leq 0, x \in [n_i, d_i], \forall 1 \leq i \leq l \), and \( v(x) \geq 0, x \in [d_i, n_{i+1}], \forall 1 \leq i \leq l - 1 \), and \( m_1 = x_N, b_n = x_D \). There is always a rational number in any interval, so the number
of intervals is countable.
\[
\int_{x_N}^{x_D} q_1(x)v(x)dx
\]
\[
= \sum_{i=1}^{l} \int_{n_i}^{d_i} q_1(x)v(x)dx + \sum_{i=1}^{l-1} \int_{n_i+1}^{d_i} q_1(x)v(x)dx
\]
\[
\leq \sum_{i=1}^{l} q_1(n_i) \int_{n_i}^{d_i} v(x)dx + \sum_{i=1}^{l-1} q_1(n_{i+1}) \int_{d_i}^{n_i+1} v(x)dx \tag{4}
\]
\[
= q_1(n_1) \int_{x_N}^{x_D} v(x)dx + \sum_{i=2}^{l} (q_1(n_i) - q_1(n_{i-1})) \int_{d_{i-1}}^{d_i} v(x)dx
\]
\[
\leq q_1(n_1) \int_{x_N}^{x_D} v(x)dx + \sum_{i=2}^{l} (q_1(n_i) - q_1(n_{i-1})) \int_{x_{i-1}}^{x_i} v(x)dx \tag{5}
\]
\[
= q_1(n_1) \int_{x_N}^{x_{i-1}} v(x)dx + (q_1(n_i) - q_1(n_1)) \int_{x_{i-1}}^{x_i} v(x)dx
\]
\[
= q_1(n_1) \int_{x_N}^{x_{i-1}} v(x)dx + q_1(n_1) \int_{x_{i-1}}^{x_i} v(x)dx
\]
\[
\leq q_1(x_N) \int_{x_N}^{x_{i-1}} v(x)dx + 1 \times \int_{x_{i-1}}^{x_i} v(x)dx
\]

The inequality (4) is based on the facts that \(q_1(x)\) is a weakly monotone increasing function and it becomes an equality by setting \(q_1(x) = q_1(x_N), x \in [x_N, x'], q_1(x) = 1, x \in [x', x_D]\). Let \(g(s) = \int_s^{x_D} v(x)dx\). In equality (5), we define \(x'\) to be the value such that \(g(x')\) achieves the maximum value in \([x_N, x_D]\).

So there are at most two pieces in \(u(\mathcal{B}D)\). \(\square\)

Similarly, we show that \(u(CN)\) is piecewise linear as well.

**Lemma 5.10.** There is an optimal utility function such that \(u(CN)\) is piecewise linear with at most 2 pieces.

**Proof.** The proof is similar to Lemma 5.3. Rewrite the revenue formula of region \(CSMN\) as follows,

\[
R_{CSMN} = \int_{CSMN} \mathcal{J} \cdot \mathbf{n} ds - \int_{CSMN} \Delta(z)u(z)dz
\]
\[ \int_{CSMN} J \cdot \mathbf{n} ds + \int_{x_C}^{x_N} y_C u(x, y_C) f_1(y_C) f_1(x) dx \]
\[- \int_{x_C}^{x_N} \int_{y_C - u(x, y_C)}^{y_C} \Delta(x, y) u(x, y) dy dx \]
\[= \int_{x_C}^{x_N} [y_C u(x, y_C) f_1(y_C) f_1(x) - \int_{y_C - u(x, y_C)}^{y_C} (u(x, y_C) - y_C + y) \Delta(x, y) dy] dx + C_2 \]
\[= \int_{x_C}^{x_N} R(u(x, y_C), x) dx + C_2 \]

We let \( C_2 = \int_{CSMN} J \cdot \mathbf{n} ds \), it only depends on \( u(C) \) and \( u(N) \). Let \( R(u(x, y_C), x) = y_C u(x, y_C) f_1(y_C) f_1(x) - \int_{y_C - u(x, y_C)}^{y_C} (u(x, y_C) - y_C + y) \Delta(x, y) dy \).

Pick point \( K \) in \( CN \), we have,
\[ \frac{\partial R}{\partial u}(u(x_K, y_C), x_K) = f_1(x_K)[y_C f_1(y_C) - \int_{y_C - u(x_K, y_C)}^{y_C} f_1(y)] dy + PR(f_1(x_K)) + PR(f_1(y))dy \]

While \( x_K \) increases, \( u(x_K, y_C) \) and \( PR(f_1(x_K)) \) increase, \( y_C f_1(y_C) - \int_{y_C - u(x_K, y_C)}^{y_C} f_1(y) \) decreases. WLOG, we can assume \( \frac{\partial R}{\partial u}(u(x, y_C), x) \geq 0, x \in [x_C, x_K] \) and \( \frac{\partial R}{\partial u}(u(x, y_C), x) \leq 0, x \in [x_K, x_N] \). The revenue increases as \( u(KN) \) decreases. The revenue increases as \( u(CK) \) increases. Let \( u^1(x, y_C) = u(C) + \frac{u(K) - u(C)}{x_K - x_C}(x - x_C), u^2(x, y_C) = u(N) + q_1(x_N, y_N)(x - x_N) \), then \( u(x, y_C) = \max(u^1(x, y_C), u^2(x, y_C)) \), \( px \in [x_C, x_N] \) gives the optimal revenue with fixed \( u(C), u(K), u(p)(N) \) and \( q_1(x_N, y_C) \). So, the optimal \( u(CN) \) comprises of at most two pieces.

With the two lemmas above, we are able to prove Theorem 5.8.

Proof of Theorem 5.8.

When \( q_1(x_N) \) are fixed, \( q_1(x \in (x_N, x_D)) \) is irrelevant to \( R_{CSMN} \). When \( u(C) \) and \( u(N) \) are fixed, \( u(x, y_B)(x \in (x_C, x_M)) \) is irrelevant to \( R_{NMGD} \).

Look at Fig. 12. When \( u(C), u(K), u(N), q_1(x_N) \) are fixed, the buyer’s utility on the red dashed line is greater than that on the black solid line between \( CK \). The buyer’s utility on the red dashed line is less than that on the black solid line between \( KN \). According to Lemma 5.9 and 5.10, \( R_{CSMN} \) and \( R_{NMGD} \) are larger, so the total revenue is larger. Furthermore, the optimal utility function on \( ND \) is piecewise linear with two slopes: \( q_1(x_N) \) and 1. To sum up, the
optimal utility on $CD$ is piecewise linear with at most three pieces. Therefore, the optimal mechanism is of the following form:

$$
\begin{array}{ccc}
q_1 & q_2 & t \\
0 & 0 & 0 \\
1 & \alpha_1 & t_1 \\
1 & \alpha_2 & t_2 \\
\alpha_1 & 1 & t_1 \\
\alpha_2 & 1 & t_2 \\
1 & 1 & t_3 \\
\end{array}
$$

**Remark 3.** Menicucci et al. (2015) also study the case where the two items are from i.i.d distributions and each density function satisfies Condition 2. The goal of Menicucci et al. (2015) is to study when the bundling is optimal, if the utility at point $C$ is fixed; while our goal is to understand the optimal utility curve when keeping utility at both points $C$ and $D$ fixed.

Furthermore, Theorem 5.8 can be extended to the correlated distribution. When Theorem 5.8 is extended to correlated distributions, Condition 5 is changed to

$$
\forall y, \frac{f(\lambda, y)}{f(\lambda, y_C)} [3 + \frac{\lambda f_x(\lambda, y)}{f(\lambda, y)} + \frac{y f_y(\lambda, y)}{f(\lambda, y)}] \text{ is weakly monotone in } \lambda
$$

Where $f_x$ and $f_y$ denote the derivatives in the $x$ and $y$ directions. The proof is exactly the same.
Condition 5 is general enough to admit a large variety of density functions.

To have a sense of what these functions are, we have the following two propositions.

**Corollary 5.11.** If \( h(x), x \in [x_A, x_B] \) is a convex, weakly monotone increasing density function, and \( x_B h'(x_B) \leq h(x_B) \), then \( h(x) \) satisfies Condition 5.

**Proof.** We prove for the case where \( h \in C^2 \), the proof for general functions is similar. We need to prove \( \left( \frac{xh'(x)}{h(x)} \right)' \geq 0 \), i.e.,

\[
(xh^{(2)}(x) + h^{(1)}(x))h(x) - xh^{(1)}(x)h^{(1)}(x) \geq 0
\]

Since \( h \) is convex, \( h^{(2)}(x) \geq 0 \), we only need to prove \( h^{(1)}(x)h(x) - xh^{(1)}(x)h^{(1)}(x) \geq 0 \), i.e. \( h(x) - xh^{(1)}(x) \geq 0 \). Notice that,

\[
(h(x) - xh^{(1)}(x))' = h^{(1)}(x) - h^{(1)}(x) - xh^{(2)}(x) = -xh^{(2)}(x) \leq 0
\]

The last inequality is because \( h^{(2)}(x) \geq 0 \). Since we have the condition \( x_B h'(x_B) \leq h(x_B) \), the proof completes. \( \square \)

It is easy to check that \( h(x) = a_n x^n, a_n \geq 0 \) also satisfies Condition 5. In fact, there are many other functions that satisfy Condition 5.

**Corollary 5.12.** If \( h_1 \) and \( h_2 \) both satisfy Condition 5, then \( h_1 + h_2 \) and \( h_1 \cdot h_2 \) both satisfy Condition 5. Particularly, for all nonnegative-coefficient polynomials \( h(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0, a_i \geq 0, i = 0, \ldots, n \), including the functions whose Taylor coefficients are nonnegative. Examples are, \((1 - x)^{-1}, e^x, e^{e^x}, \ln \frac{1}{1-x}, \tan x (|x| < \frac{\pi}{2}), \text{hyperbolic sine} \text{ sinh } x = \frac{e^x - e^{-x}}{2}, \text{ and } \frac{F_n x}{1-(F_{n-1}+F_{n+1})x-(-1)^n x^2} \text{ where } F_n \text{ denotes the Fibonacci numbers.}

**Proof.** We prove for the case where \( h_i \in C^2, i = 1, 2 \), the general case is similar. Since \( \frac{xh_i'(x)}{h_i(x)} \) is weakly monotonically increasing, we get \( \left( \frac{xh_i'(x)}{h_i(x)} \right)' \geq 0 \), i.e.,

\[
(xh_i^{(2)}(x) + h_i^{(1)}(x))h_i(x) \geq x(h_i^{(1)}(x))^2, i = 1, 2
\]
In order to show that $h_1 + h_2$ still has monotone increasing power rate, it is equivalent to show:

$$\left(x \frac{h_1'(x) + h_2'(x)}{h_1(x) + h_2(x)}\right)' \geq 0,$$

i.e.,

$$x \left(h_1'(x) + h_2'(x)\right) \frac{h_2(x)}{h_1(x)} + x \left(h_1'(x)\right)^2 \frac{h_1'(x)}{h_2(x)} \geq 2h_1'(x)h_2'(x)$$

The proof for the case $h_1 + h_2$ is complete.

In order to show $h_1h_2$ has monotone increasing power rate, it is equivalent to show:

$$\left(x \frac{h_1'(x)h_2(x) + h_1(x)h_2'(x)}{h_1(x)h_2(x)}\right)' \geq 0,$$

i.e.,

$$x \left(h_1'(x)h_2(x) + h_1(x)h_2'(x)\right) \frac{h_2'(x)}{h_1'(x)} + x \left(h_1'(x)h_2'(x)\right) \geq 2h_1'(x)h_2'(x)$$

which is straightforward. The proof for the case $h_1h_2$ is complete.

5.4. Optimal mechanisms for uniform distribution under unit-demand constraint

A buyer is unit-demand if $q_1(x, y) + q_2(x, y) \leq 1$. Under the unit-demand model, Pavlov (2011b, Proposition 2) states that, if the distribution functions satisfy Condition 2, it is without loss of generality to restrict our attention to mechanisms such that

$$q_1(x, y) + q_2(x, y) \in \{0, 1\} \quad \forall (x, y)$$

Pavlov solves the optimal mechanism problem for two items with identical uniform distributions. The resulting mechanism contains 5 menu items for uniform distribution on $[c, c + 1] \times [c, c + 1], c \in (1, \bar{c})$ (where $\bar{c} \approx 1.372$). We show that in nonidentical settings, the optimal mechanism also contains at most 5 menu items. It follows trivially that our result is tight.

**Theorem 5.13.** In the unit-demand model, if both $f_1$ and $f_2$ are uniform distributions, there is an optimal mechanism that consists of at most 5 menu items.

Let $ASE$ denote the zero utility region and $CSEBD$ the non-zero utility region. For the same reason as in Lemma 5.1, $ASE$ is convex. For points in $ASE$, allocation $(0, 0)$ is the best. For $(x, y) \in CSEBD, (q_1(x, y), q_2(x, y)) \neq (0, 0)$, allocation $(0, 0)$ is the best.
Figure 13: Optimal unit demand allocation.

(0, 0), so \( q_1(x, y) + q_2(x, y) = 1 \). The mechanism is shown in Fig. 13. Draw a 45 degree line across \( E \), intersecting \( BD \) or \( CD \) at \( W \). Draw a 45 degree line across \( S \), intersecting \( BD \) or \( CD \) at \( G \). We consider here the case where \( W \) is on \( BD \) and \( G \) is on \( CD \). The other cases follow from similar arguments.

The theorem can be similarly proved via the following two lemmas.

**Lemma 5.14.** There is an optimal utility function such that \( u(BW) \) is piecewise linear with at most 2 pieces.

**Proof.**

For \( (x, y) \) in \( CSEBD \), the utility of choosing \((q, 1 - q)\) is \( xq + y(1 - q) - t(q) = (x - y)q + y - t(q) \). The buyer will choose the menu item that achieves \( \max_q\{(x - y)q + y - t(q)\} = y + \max_q\{(x - y)q - t(q)\} \). This means that which menu item will be chosen depends entirely on the value of \( x - y \). For two points \((y_1 + l, y_1)\) and \((y_2 + l, y_2)\) in \( CSEBD \), they have the same favorite menu items. We can assume, WLOG, that they choose the menu item with the highest payment (Hart and Reny, 2012).

Let the allocation rule on point \((x, y)\) in \( MDBE \) be \((1 - q_2(y), q_2(y))\). Fixing \( u(W) \) and \( q_2(y_W) \), for any point \((x, y)\) in region \( EBW \), the utility of the buyer is

\[
u(x, y) = u(W) - (u(W) - u(x_B, y + x_B - x)) - (u(x_B, y + x_B - x) - u(x, y))
\]
\[ R_{EBW} = \int_{EB} \mathcal{J} \cdot \hat{n} ds - \int_{EBW} \Delta(z) u(z) dz = \int_{WE} \mathcal{J} \cdot \hat{n} ds + \int_{BW} \mathcal{J} \cdot \hat{n} ds + \int_{EB} \mathcal{J} \cdot \hat{n} ds - \int_{EBW} \Delta(z) u(z) dz \] (6)

The second term in the expression above is

\[
\int_{BW} \mathcal{J} \cdot \hat{n} ds = \int_{yb}^{yw} x_B (x_B - x_E) f_1 f_2 dy + \int_{yb}^{yw} x_B q_2(y) (y - y_B) x_B f_1 f_2 dy + x_B (x_B - x_E) (y_W - y_B) f_1 f_2
\]

where \( v_1(y) = (y_B - y) x_B f_1 f_2 \) and \( C_1 = x_B (x_B - x_E) (y_W - y_B) f_1 f_2 \). \( v_1 \) and \( C_1 \) are independent of \( q_2(y), y \in [y_B, y_W] \). Since in this part, we are only concerned with the form of the optimal \( q_2(y) \), we use this shorthand notation for simplicity.

Similarly, for the rest of the terms in Equation (6), we have

\[
\int_{EB} \mathcal{J} \cdot \hat{n} ds = \int_{yc}^{yw} q_2(y) v_2(y) dy + C_2
\]
\[
\int_{EW} \mathcal{J} \cdot \hat{n} ds = \int_{yc}^{yw} q_2(y) v_3(y) dy + C_3
\]
\[
\int_{EBW} \Delta(z) u(z) dz = \int_{yc}^{yw} q_2(y) v_4(y) dy + C_4
\]
where, $v_i$ and $C_i$, $i = 2, 3, 4$ are independent of $q_2(y), y \in [y_B, y_W]$. Then, Equation (6) becomes

$$\int_{y_C}^{y_W} q_2(y)(v_1(y) + v_2(y) + v_3(y) - v_4(y))dy + C_1 + C_2 + C_3 - C_4$$

The following proof is similar to Lemma 5.9; the optimal $q_2(y), y \in [y_C, y_W]$ comprises two parts. There is $y' \in [y_C, y_W]$ such that

$$q_2(y) = \begin{cases} 0 & y \in [y_C, y') \\ q_2(y_W) & y \in [y', y_W] \end{cases}$$

The utility function is

$$u(x_B, y) = \begin{cases} u(W) + (y - y_W)q_2(y_W) & y \in [y', y_W] \\ u(W) + (y' - y_W)q_2(y_W) & y \in [y_B, y'] \end{cases}$$

Hence $u(x_B, y) \geq u(W) + (y' - y_W)q_2(y_W) \geq u(W) + (y_B - y_W)$. Because $WE$ is a 45 degree line, $x_B - x_E = y_W - y_B$. Then $u(W) = (y_W - y_B)q_2(y_W) + (x_B - x_E)(1 - q_2(y_W)) = y_W - y_B$. We get $u(x_B, y) \geq 0$. This new utility function satisfies the convexity and nonnegativity property, so it is feasible. Therefore, $u(BW)$ comprises of at most two pieces.

**Lemma 5.15.** There is an optimal utility function such that $u(WD)$ is piece-wise linear with at most 2 pieces.

**Proof.**

$$R_{GSMEW} = \int_{x_G}^{x_D} J \cdot \hat{u} ds + \int_{y_C}^{y_W} J \cdot \hat{u} ds + \int_{y_C}^{y_W} J \cdot \hat{u} ds - \int_{GSM} 3fu(z)dz - \int_{DM} 3fu(z)dz$$

$$= C + \int_{x_G}^{x_D} y_C u(x, y_C)dx - 3\int_{x_G}^{x_D} \int_{y_C}^{y_W} (y - y_C + u(x, y_C))dydx$$

$$+ \int_{y_W}^{y_D} x_B u(x_B, y)dy - 3\int_{y_W}^{y_D} \int_{x_W}^{x_D} (x - x_B + u(x, y))dxdy$$

$$= C + \int_{x_G}^{x_D} fu(x, y_C)(y_C - \frac{3}{2}u(x, y_C))dx + \int_{y_W}^{y_D} fu(x_B, y)(x_B - \frac{3}{2}u(x_B, y))dy$$

So, for fixed $u(D)$ and $u(W)$, when $u(x_B, y) > \frac{1}{3}x_B, x \in (y_W, y_D)$, the revenue is decreasing in $u(x_B, y)$; when $u(x_B, y) < \frac{1}{3}x_B, y \in (y_W, y_D)$, the revenue is
increasing in \( u(x_B, y) \). Similar to the proof in Lemma 5.3, wlog, we can assume that there is a point \( P \) on the \( WD \) segment, such that \( u(x_B, y_P) = \frac{1}{3} x_B \).

Since the utility at any point in \( PD \) is equal to or larger than \( u(P) \), the revenue increases as \( u(PD) \) decrease. Since \( u(WP) \leq u(PD) \), the revenue increases as \( u(WP) \) increases. Let \( u_1(x_B, y) = u(W) + u(PD) - u(W) \), \( u_2(x_B, y) = u(D) + q_2(y_D)(y - y_D) \). Then \( u(x_B, y) = \max\{u_1(x_B, y), u_2(x_B, y)\} \). So, the optimal \( u(WD) \) has at most two pieces.

\[ \square \]

Proof of Theorem 5.13.

Figure 14: optimal utility function.

We can now settle the optimal utilities on \( BD \), subject to fixed values of \( u(W) \) and \( u(P) \) as well as the convexity of \( u \). Let \( \alpha = \frac{u(P) - u(W)}{y_P - y_W} \) and \( t_\alpha = (1 - \alpha)x_P + \alpha y_P - u(P) \). Adding a new menu item \( (1 - \alpha, \alpha, t_\alpha) \), the utility of choosing this new item is \( u^\alpha \), which is denoted by the red dashed line shown in Fig. 14. Let \( u^D \) be the utility obtained by choosing the same menu item as point \( D \): \( (1 - q_2(y_C), q_2(y_C), t^D) \). Thus, for any point \( (x, y_C) \) on \( CD \), we have

\[
\begin{align*}
    u(x, y_C) &\geq u^D(x, y_C) = (1 - q_2(y_C))x + q_2(y_C)y_C - t^D \\
    &\geq (1 - q_2(y_C))x_D + (1 - q_2(y_C))(x - x_D) + q_2(y_C) - t^D \\
    &= u(D) + (1 - q_2(y_C))(x - x_D) \geq u^\alpha(D) + (1 - \alpha)(x - x_D) \\
    &= (1 - \alpha)x_D + \alpha y_C - t^\alpha + (1 - \alpha)(x - x_D) = u^\alpha(x, y_C)
\end{align*}
\]
Thus, after adding this new menu item, $u(CD)$ does not change and $u(PW)$ will weakly increase.

In the optimal mechanism with fixed $q_2(y_W)$, according to Lemma 5.14, the types on $BW$ choose only two menu items: $(1 - \alpha, \alpha, t_\alpha)$ and $(1, 0, t_1)$ for some $t_1$. Using the same arguments as for $CD$, for $BD$, we have another two menu items: $(1 - \beta, \beta, t_\beta)$ and $(0, 1, t_2)$ for some $t_2$.

On edge $BD$, we pick two points $W$ and $P$. Using the same method, we pick two points $G$ and $O$ on edge $CD$ correspondingly. We have proved that adding the new menu item $(1 - \alpha, \alpha, t_\alpha)$ does not change $u(CD)$ and $u(PD)$. Similarly, adding the new menu item $(1 - \beta, \beta, t_\beta)$ does not change $u(BD)$ and $u(OD)$. We now consider a brand new menu with only the above four menu items and $(0, 0, 0)$. Compared to the previous menu, $u(PW)$ and $u(GO)$ weakly increase. $u(PD)$ and $u(OD)$ weakly decrease since there are less menu items.

Following a proof similar to Lemma 5.15, $R_{GSM EW D}$ weakly increases. For fixed $q_2(y_W)$ and $u(W)$, $R_{EB W}$ is maximized. For fixed $q_1(x_G, y_G)$ and $u(G)$, $R_{C SG}$ is maximized. The total revenue weakly increase. The optimal menu consists of at most 5 menu items:

<table>
<thead>
<tr>
<th>$q_1$</th>
<th>$q_2$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$t_1$</td>
</tr>
<tr>
<td>$1 - \alpha$</td>
<td>$\alpha$</td>
<td>$t_\alpha$</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>$t_2$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$1 - \beta$</td>
<td>$t_\beta$</td>
</tr>
</tbody>
</table>

Haghpanah and Hartline (2015) study the same setting as ours. Compared to their paper, in this section, we focus on the uniform distribution and characterize the optimal mechanism, while they proceed in the reverse direction, by first assuming that the optimal mechanism must be uniform pricing and then infer the distribution conditions that imply the optimality of uniform pricing. The distributions derived by their approach satisfy a certain monotonicity condition and are not limited to the uniform distribution.
For this part of our work, it may be difficult to generalize our technique to high dimensional settings. Let us consider the case of selling three items. In this case, the buyer’s valuation is a cube. Pavlov’s characterization is still correct in that the seller either keeps all the items, or sells one of the items with probability 1. Then, we must figure out what the allocation probability of the other two items is. There are two difficulties generalizing the current technique. First, our technique is to fix the utility on 5 surfaces and optimize the utility on the remaining surface. However it is a two dimensional optimization problem, and there is no valid technique to solve it. On the other hand, even if we restrict to the deterministic allocation, there are eight allocation probabilities to consider. It is hard to derive a result that the optimal mechanism consists of only a few menu items.

The technique cannot be extended to environments when the seller has independent positive costs on the goods. It is equivalent to considering that the seller has zero costs on the goods and the buyer has negative valuation of the goods. In that case, $x_A < 0$ or $y_A < 0$, which brings up many new cases to consider. The current technique cannot directly generalize to this case either.

6. Additional Related Work

When the distributions are discrete, a recent series of papers (Daskalakis and Weinberg, 2011; Cai et al., 2012a,b; Alaei et al., 2012) show that the general optimal mechanism ($k > 1$) is the solution of a linear program. They provide different algorithmic methods to solve the linear program. For continuous distributions, Both Chawla et al. (2010) and Cai and Huang (2013) study the possibility of designing simple auctions that achieve approximately optimal revenue in the worst case. In addition, (Daskalakis and Weinberg, 2012; Cai and Huang, 2013) provide a polynomial time approximation scheme (PTAS)$^9$ of the optimal auction under assumptions on the distributions.

---

$^9$In computer science, a PTAS is an approximate algorithm that produces a solution within a factor $1 + \epsilon$ of the optimal solution in polynomial time.
Hart and Nisan further extend the approximation results on selling two independent items to the general case with \( k \) independent items: separate sale guarantees at least a \( \frac{c}{\log k} \)-fraction of the optimal revenue; for identically distributed items, bundling guarantees at least a \( \frac{c}{\log k} \)-fraction of the optimal revenue. Li and Yao (2013) tighten these lower bounds to \( \frac{c}{\log k} \) and \( c \) respectively. Furthermore, among several other results, Yao (2015) shows that a simple auction, namely the better of the Vickrey auction with entry fees and Ronen’s lookahead auction (Ronen, 2001), can always yield a constant fraction of the optimal revenue in the general multi-dimensional setting where the joint type distributions are independent among items and can be correlated among buyers.

7. Acknowledgement

The authors would like to acknowledge Aris Filos Ratsikas for careful proofreading. The authors also have benefited by discussions with Gregory Pavlov, Jason Hartline and participants of the second workshop of “New trends in mechanism design II”. This work was supported in part by the National Natural Science Foundation of China Grant 61303077, 61561146398, a Tsinghua University Initiative Scientific Research Grant and a China Youth 1000-talent program.

References


54