Overview: A Simple and Approximately Optimal Mechanism for an Additive Buyer

Abstract

In this report, we present an overview of the paper *A Simple and Approximately Optimal Mechanism for an Additive Buyer* [1]. We focus on multi-parameter mechanism design and use a counter-intuitive example based on what we have learned in class to imply the richness of the problem, as well as providing motivation for further results. Then, we analyze primary contributions of the paper and apply life experiences to understanding the main result by intuition. Through previous steps, we can have a clear understanding of the core idea. That is, the mechanism of using the better of selling items separately and bundling together can a achieve a constant-factor approximation to the optimal revenue. Then, we go deep into the formal result and its proof. To make the proof easier to understand for beginners in AGT like us, we add in more details for the confusing part and reconstruct the whole proof to clarify the logic behind. The main proof is presented in two parts: bounding the revenue from the tail and bounding the welfare from the core. Eventually, we discuss extensions to multiple buyers and show this method fail to have a constant-factor approximation in the multiple-buyer case, which we believe is its biggest limitation.

Contents

1 Problem statement

1.1 The setting: multi-parameter mechanism design

In this paper, we focus on **multi-parameter mechanism design**. We start from the simplest version of this problem, and then generalize it to more complex situations.

> A monopolist seller owns n heterogeneous items with no value for them, and a single buyer has non-negative value v_i for each item, the exact number of which the seller doesn't know, but knows that each of v_i is sampled independently from a distribution D_i.

In this case, The seller aims to **maximize the expected revenue** by setting payment rules and the buyer aims to maximize her expected utility when deciding private values.

1.2 A simple yet counter-intuitive example

We consider the case of an additive buyer, whose value of a set of S items is $\sum_{i \in S} v_i$. In this case, it seems that there's no interaction between items from the buyer's perspective: the buyer's value for item i does not depend on other items and the seller only knows the independent distribution of values for different items. Intuitively, we may treat this as n **separate single-parameter mechanism design**. Thus, recalling what we have learned in class (Myerson's theory), the expected revenue for a price p_i

$$p_i \cdot (1 - F_i(p_i)) \tag{1.1}$$

where $F_i(p_i)$ refers to the probability that a random variable drawn form F_i has value at most p_i. Set the price p_i that makes the above revenue highest, and we can know the maximum expected revenue. That's it?

Unfortunately, this mechanism is not optimal, even for the following simple case[2]:

Example. Consider the case of two items, where the buyer's value for each item is drawn i.i.d from the uniform distribution over the finite set $\{1, 2\}$. That is,

$$\text{Two goods } X_1, X_2, \text{ independent}$$
$$X_1, X_2 = \begin{cases} 1, & \text{with probability } \frac{1}{2} \\ 2, & \text{with probability } \frac{1}{2} \end{cases}$$

1.2.1 Selling separately

Treating each item as a separate problem, we may use Myerson's theory as mentioned above. For each item, we can either choose to price the item at 1, selling with absolute probability and obtaining a revenue of 1, or choose to price it at 2, selling with probability $1/2$ and again getting an expected revenue of 1. In this way, the seller can obtain expected revenue one per item and two in total. The calculation process is as follows:

$$\begin{aligned} \text{REV}(X_1) &= \max\{1 \cdot (1 - 0), 2 \cdot (1 - \frac{1}{2})\} = 1 \\ \text{REV}(X_2) &= \max\{1 \cdot (1 - 0), 2 \cdot (1 - \frac{1}{2})\} = 1 \\ \text{REV}(X_1) &+ \text{REV}(X_2) = 2 \end{aligned}$$

1.2.2 Bundling together

However, there's another mechanism that offers the buyer only two choices:

1. receive *both* items together for 3 (called bundling together)
2. receive nothing for 0

For choice 1, the possible price for two items is $p_{12} = \{2, 3, 3, 4\}$. Hence, the probability that p_{12} is greater than 3 is $1/4$. Using Myerson's theory to treat the bundle as a whole, we can get a higher expected revenue

$$\mathrm{REV}(X_1) + \mathrm{REV}(X_2) = 3 \cdot \left(1 - \frac{1}{4}\right) = \frac{9}{4} > 2$$

Thus, it is suboptimal to sell each item separately, which is counter-intuitive. In this case, bundling together works better. However, there are other cases when bundling together doesn't work either. We must therefore conclude that

> Neither selling separately nor bundling together guarantees a constant fraction of the optimal revenue.

Despite the conclusion, the two methods provide motivation for a constant-factor approximation to the optimal revenue, which will be presented in chapter 2 and 3.

1.3 What are the challenges?

From the example above, it is clear that optimal mechanisms are richer than we have expected. Still, we may reasonably hope that optimal mechanisms are not too complex. Unfortunately, this is not the case, and prior work has identified three primary challenges for the problem:

- Menu complexity: the number of distinct options available for the buyer to purchase is uncountable.
- Computational intractability: it is #P-hard to find the optimal auction in certain settings.
- Non-monotonicity: the optimal revenue for a "strictly better" distribution is strictly worse.

In contrast, the method this paper adopt has solved these challenges successfully, with polynomial menu complexity, being revenue-monotone, and implementable in poly-time.

2

Core idea of the paper

2.1 Primary contributions

To get the core idea of the paper, we need to grasp primary contributions macroscopically:

- Main result: a **constant-factor approximation** to the optimal revenue guaranteed by means of *a priori maximum* of the revenue generated by selling separately and bundling together.
- Practical value: useful in practice (as multi-item setting is incredibly normal in real life) and is the **FIRST** computationally tractable mechanism to obtain a constant-factor approximation for simple mechanisms.
- Extensions: to multiple buyers and to valuations that are correlated across items.

2.2 Intuitive presentation of the result: by real-life examples

To begin with, an informal presentation of the main result is as follows:

Main Result (Informal). In any market with a single additive buyer with independent item values, either selling separately or bundling together guarantees a 6-approximation to the optimal revenue.

Next, let me give you a sense of why the main result holds by following intuitive (yet maybe inappropriate) examples:

1. **Bundling together**: when the sum of the buyer's values for all items tends to **concentrate** around its expectation. For example, in the grocery store where a dozen of yogurt costs 15 yuan and two bags of milk costs 10 yuan, the sum concentrates and thus items are bundled together (in practical terms) at a price of 22 yuan. As the milk may be approaching its expiry date and there's high possibility that it will not be sold before expiration, bundling the two together at a price slightly below the expected value will extract a significant fraction of the total value as revenue.

2. **Selling separately**: when the sum is dominated by rare events where one or more items' value is significantly higher than usual. The existence of these items is called "**tail**" event in the paper. For example, in a clothing store, a hand-made suit is priced at 100000 yuan and a fashionable hat is priced (relatively) much lower at 100 yuan. The hat is quite popular around people and many of them are willing to queue up to buy. If we bundle the hat and the tail (namely the suit) together at 100000 yuan, due to the high price of the suit, people interested in buying the hat will plummet, which will result in the sharp drop of expected revenue. Hence, in this case, the best choice is to sell these things separately and to optimize the revenue from each individual item by itself.

Actually, selling things separately and bundling things together are both commonly seen and widely used in a variety of real-life situations, let alone the idealized auction. Sometimes selling things separately works and the other doesn't, sometimes conversely (see the two examples above). Thus, with our common sense of life, we can have the intuition that making a choice between the two (as most of sellers do) will guarantee an approximation to the optimal revenue.

However, the boundary between selling separately and bundling together is yet unclear. To take an extreme example, it can be described as concentrated even when each of the item has an expected value of infinity. For n i.i.d. Equal Revenue distributions, the expected value for each item is infinite, yet the expected optimal value is finite. In this case, bundling together is approximately optimal for the distribution and the distribution of the sum of items concentrate. Thus, Yi and Yao have developed a **Core-Tail Decomposition method** [3] which proposes adequate definitions for tail events, and separate the tail event (called "the tail" afterwards) from the rest (called "the core" afterwards).

3

Formal proof of the result

In this chapter, we will give a proof of the main result clarified informally last chapter. In the first place, we state the formal description of it.

For $m = 1$ buyer and n items, $\text{REV}(D) \leq 6 \max\{\text{SREV}(D) + \text{BREV}(D)\}$.

where

- $\text{REV}(D)$: The optimal revenue obtained by any Bayesian Incentive Compatible (**BIC**) or Ex-Post Individually Rational (**IR**) mechanism when the buyer profile is drawn from D.
- $\text{SREV}(D)$: The optimal revenue obtained by **selling items separately** when the buyer profile is drawn from D. In practice, we just need to run Myerson's optimal auction separately for each item. In this case where we only have $m = 1$ buyer, $\text{SREV}(D)$ is achieved by setting a price p_i on each item i, and letting the buyer pick any subset of items to purchase.
- $\text{BREV}(D)$: The optimal revenue obtained by **selling the grand bundle** when the buyer profile is drawn from D. For $m = 1$ buyer, $\text{BREV}(D)$ is achieved by setting a p , and letting the buyer purchase the grand bundle if she is willing to pay.

3.1 Preliminaries

3.1.1 Notation definition: for convenience

For convenience, we will first list notations alphabetically before starting the proof (as I was so easily drowned in the sea of notations when reading and had to turn the paper forward and backward constantly to find out their meanings). You can keep this page open in your own way as you read the subsequent part for better understanding. Some settings are defined more generally for the sake of further extensions.

Notation	Definition
A	A subset of items, often as the items whose values lie in the tail of their respective distributions.
D	The joint $m \cdot n$-dimensional distribution over all buyers for all items.
D_i	The m-dimensional distribution over all buyers' values for item i.
D^j	The n-dimensional distribution over buyer j's values for all items.
D_i^C	The core of D_i, the conditional distribution of $\vec{v}_i$ conditioned on $v_i^{\leq t_i r_i}$.
D_i^T	The tail of D_i, the conditional distribution of $\vec{v}_i$ conditioned on $v_i^{>t_i r_i}$.
D_A^T	The product distribution equal to $\times_{i \in A} D_i^T$.
D_A^C	The product distribution equal to $\times_{i \notin A} D_i^C$, representing the distribution of values in the core, conditional on A being the set whose values lie in the tail.
D_A	$D_A^C \times D_A^T$. That is, D_A is the distribution D, conditioned on $v_i^* > t_i r_i$ for all $i \in A$ and conditioned on $v_i^* \leq t_i r_i$ for all $i \notin A$.
p_i	$Pr[v_i^* > t_i r_i]$. The probability that the highest value on item i lies in the tail, depending both on the distribution D_i and the choice of t_i.
p_A	$(\prod_{i \in A} p_i)(\prod_{i \notin A}(1 - p_i))$. Equals the probability $Pr[\vec{v} \in \text{support}(D_A)]$ when it is not null.
r_i	The optimal revenue obtainable by selling just item i.
r	$\sum_i r_i$. The total revenue from optimally selling the items separately. Equals $\text{SREV}(D)$ and introduced for convenience.
t_i	A parameter for item i to define the separation between the core and tail of distribution D_i. That is, the core for item i will be supported on the interval $[0, t_i r_i]$, and the tail for item i will be supported on $[t_i r_i, \infty]$.
$\vec{v}$	A random sample from D.
$\vec{v}_i$	A random sample from D_i.
$\vec{v^j}$	A random sample from D^j.
v_i^*	The maximum value for item i.

To act as the bridge of the proof, we also introduce the concept of welfare.

- $\text{VAL}(D)$: $E_{\vec{v}\sim D}[\sum_i v_i^*]$. The expected optimal welfare.
- $var(D)$: $\text{var}_{\vec{v}\sim D}(\sum_i v_i^*)$. Variance of the welfare.

3.1.2 Lemma statement: for proof

We will list all lemmas at once and explain their meanings later. For simplification, no proofs are given except for Lemma 6. See references for more details.

- **Lemma 1.**[2, 4] $\text{REV}(D \times D') \leq \text{VAL}(D) + \text{REV}(D')$, where D and D' are distributions over values for disjoint sets of items.
- **Lemma 2.** [2] $\text{REV}(D) \leq n \cdot m \cdot \text{SREV}(D)$.
- **Lemma 3.** [3] $p_i \leq 1/t_i$ for all i.
- **Lemma 4.** [3] $\text{REV}\left(D_i^C\right) \leq r_i$ and $\text{REV}\left(D_i^T\right) \leq r_i/p_i$
- **Lemma 5.** [2] $\text{REV}\left(D\right) \leq \sum_A p_A \text{REV}\left(D_A\right)$.
- **Lemma 6.** (Core-Tail Decomposition) $\text{REV}(D) \leq \text{VAL}\left(D_\emptyset^C\right) + \sum_A p_A \text{REV}\left(D_A^T\right)$
- **Lemma 7.** [3] Let F be a one-dimensional distribution with optimal revenue at most y supported on $[0, ty]$. Then $var(F) \leq (2t-1)y^2$.

To explain their meanings:

1. Lemma 1 and 2: some useful bounds on $\text{REV}(D)$.
2. Lemma 3, 4 and 5: some known results about the core.
3. Lemma 6: main decomposition lemma. It relates the the optimal revenue from a distribution D to the revenue and welfare that can be extracted from the tail and core of D, respectively. That is, the optimal revenue from D is at most the **optimal revenue** from the tail plus the **expected welfare** (rather than the optimal revenue, as it is problematic for multiple buyers) from the core. It's a weaker bound, as $\text{REV}(D) \leq \text{VAL}(D)$ for all D from Individual Rationality.
4. Lemma 7: an upper bound on the variance of welfare using a small range of the core.

Now, we will give a proof of Lemma 6 in our own words using other given lemmas.
Proof: From Lemma 1, whose form is closet to Lemma 6, we set $D = D_A^C, D' = D_A^T$,

$$\mathrm{REV}(D_A) = \mathrm{REV}(D_A^C \times D_A^T) \leq \mathrm{VAL}(D_A^C) + \mathrm{REV}(D_A^T) \tag{3.1}$$

To turn $\mathrm{REV}(D_A)$ into $\mathrm{REV}(D)$, we can refer to Lemma 5, which relates the joint distribution D with the distribution of core and tail. Thus,

$$\begin{aligned}
\mathrm{REV}(D) &\leq \sum_A p_A \mathrm{REV}(D_A) \\
&\leq \sum_A p_A (\mathrm{VAL}(D_A^C) + \mathrm{REV}(D_A^T)) \\
&\leq \left(\sum_A p_A\right) \mathrm{VAL}(D_A^C) + \sum_A p_A \mathrm{REV}(D_A^T)
\end{aligned} \tag{3.2}$$

Now, we are close to the desired results. According to the definition of $\mathrm{VAL}(D_A^C)$ (can be found in the notation list), it refers to the expected sum of values for items not in the tail, which is less than the expected sum for all items $D_\emptyset^C$. That is,

$$\mathrm{VAL}(D_A^C) \leq \mathrm{VAL}(D_\emptyset^C) \tag{3.3}$$

Moreover, the sum of probability p_i equals 1 by definition. Hence,

$$\begin{aligned}
\mathrm{REV}(D) &\leq \left(\sum_A p_A\right) \mathrm{VAL}(D_A^C) + \sum_A p_A \mathrm{REV}(D_A^T) \\
&\leq \mathrm{VAL}(D_\emptyset^C) + \sum_A p_A \mathrm{REV}(D_A^T)
\end{aligned} \tag{3.4}$$

□

3.2 Main proof

Because we use "core" and "tail" to separate items and use t_i to decide the boundary between them, our proof will be mainly based on the application of Core-Tail Decomposition method mentioned above (Lemma 6). To get the desired result, we need to bound **the revenue from the tail** and **welfare from the core** separately. Hence, our proof will be presented in two parts.

As for the boundary, we let t_i scale inverse proportionally to r_i, so that high-revenue items are more likely to occur in the tail. Thus, it allows us to capture scenarios in which

- revenue comes primarily from one heavy item by analyzing the tail
- instances driven by the combined contribution of many light items by analyzing the core

If we set $t_i = cr/r_i$, then the boundary between core and tail becomes $t_i r_i = cr = c\text{SREV}(D)$. Hence, while the choice of t_i is non-uniform, the absolute cutoffs $t_i r_i$ are uniform. It turns out that this is precisely the threshold that we need for t_i for the proof of both core and tail.

To make the proof easier to understand, we add in more details for the confusing part and reconstruct the whole proof to clarify the logic behind.

3.2.1 Part one: bounding the revenue from the tail

Proposition 1. For a single buyer, when $t_i = r/r_i$ for every i, $\sum_A p_A \text{REV}\left(D_A^T\right) \leq 2\text{SREV}(D)$.

Proof: To build the relationship between $\text{REV}\left(D_A^T\right)$and $\text{SREV}(D)$, using Lemma 2 and Lemma 4, we can know that

$$\begin{aligned} \text{REV}\left(D_A^T\right) &\leq |A|\text{SREV}\left(D_A^T\right) \\ \text{REV}\left(D_A^T\right) &\leq \sum_{i \in A} |A| r_i / p_i \end{aligned} \tag{3.5}$$

According to the definition of $\text{SREV}\left(D_A^T\right)$, we can obtain that $\text{SREV}\left(D_A^T\right) = \sum_{i \in A} r_i$. As p_i symbolizes probability, $p_i \in [0, 1]$, and equation (3.5) can thus be written as

$$\text{REV}\left(D_A^T\right) \leq |A|\text{SREV}\left(D_A^T\right) \leq \sum_{i \in A} |A| r_i / p_i \tag{3.6}$$

Using what we have obtained above, we may sum $\text{REV}\left(D_A^T\right)$ to get the desire term of $\sum_A p_A \text{REV}\left(D_A^T\right)$. Then, rewrite the sum by first summing over item i, and then summing over every set A containing i. That is,

$$\sum_A p_A \text{REV}\left(D_A^T\right) \leq \sum_A p_A \sum_{i \in A} |A| r_i / p_i = \sum_i r_i \sum_{A \ni i} |A| \cdot p_A / p_i \tag{3.7}$$

In the following part, we will keep modifying the right-hand side of the equation to for the eventual form of $\text{SREV}(D)$.

Then, we may wish to interpret the term $\sum_{A \ni i} |A| \cdot p_A / p_i$. Based on the definition of p_A and p_i, we can observe that p_A / p_i is the exactly the conditional probability that the set A of items are in the tail and all other items are not (namely p_A), conditioned on i being in the tail (namely p_i). As $|A|$ is the size of A, summing over all $A \ni i$ therefore yields the expected size of the set of items in the tail, conditioned on i being in the tail. Thus, $\sum_{A \ni i} |A| \cdot p_A / p_i$ is the same as $1 + \sum_{j \neq i} p_j$ by its meaning.

Next, for the above p_j, we can further apply Lemma 3 to it, thus

$$1 + \sum_{j \neq i} p_j \leq 1 + \sum_{j \neq i} 1/t_j \tag{3.8}$$

For the above t_j, as we choose $t_j = r/r_j$ for each item and $\sum_j r_j = r$ by definition, we can obtain that

$$\sum_{A \ni i} |A| \cdot p_A / p_i = 1 + \sum_{j \neq i} p_j \leq 1 + \sum_{j \neq i} 1/t_j \leq 2 \tag{3.9}$$

Therefore, combining equation (3.7) and (3.9), we can draw a conclusion that

$$\sum_A p_A \mathrm{REV}\left(D_A^T\right) \leq \sum_i r_i \sum_{A \ni i} |A| \cdot p_A / p_i \leq 2 \sum_i r_i = 2\mathrm{SREV}\left(D_A^T\right) \tag{3.10}$$

□

3.2.2 Part two: bounding the welfare from the core

Proposition 2. For a single buyer, when $t_i = r/r_i$ for each i, it holds that $\max\{\mathrm{SREV}(D) + \mathrm{BREV}(D)\} \geq \frac{1}{4}\mathrm{VAL}(D_\emptyset^C)$.

Now, we need to handle the welfare from the core. With the use of an upper bound on the **variance** of the welfare, we can conclude that the welfare is concentrated whenever it is sufficiently large relative to $\mathrm{SREV}(D)$. Thus, within the core,

1. if the welfare is "small" compared to $\mathrm{SREV}(D)$, then selling separately extracts most of the welfare;

2. if the welfare "concentrates", then bundling together extracts most of the welfare.

Proof: We will use the variance and expectation of $D_\emptyset^C$, divide the case by its expectation and prove the proposition case by case.

case 1: $\text{VAL}\left(D_{\emptyset}^{C}\right) < 4r$

In this case, $\text{VAL}\left(D_{\emptyset}^{C}\right) < 4r = 4\text{SREV}(D)$. Hence, we can easily get $\text{SREV}(D) < \frac{1}{4}\text{VAL}\left(D_{\emptyset}^{C}\right)$.

case 2: $\text{VAL}\left(D_{\emptyset}^{C}\right) \geq 4r$

Firstly, we will introduce the upper bound on the variance as mentioned above. Using Lemma 7, since $\text{REV}\left(D_i^C\right) \leq r_i$, and the distribution D_i^C is supported on $[0, t_i r_i]$, we can know that

$$var(D_i^C) \leq 2t_i r_i^2 \tag{3.11}$$

Then, we sum over all i and recall that $t_i = r/r_i$, getting the exact upper bound for $D_{\emptyset}^{C}$

$$var(D_{\emptyset}^{C}) = \sum_i var(D_i^C) \leq 2\sum_i t_i r_i^2 = 2r^2 \tag{3.12}$$

Thus, we have $var(D_{\emptyset}^{C}) = 2r^2$ and its expectation $\text{VAL}\left(D_{\emptyset}^{C}\right) \geq 4r$. With the information above, we can recall Chebyshev's inequality,

$$\begin{aligned} Pr_{\vec{v}\leftarrow D}[\sum_i v_i \leq \frac{2}{5}\cdot \text{VAL}(D_{\emptyset}^{C})] &\leq \frac{var(D_{\emptyset}^{C})}{(1-\frac{2}{5})^2 \cdot \text{VAL}(D_{\emptyset}^{C})^2} \\ &\leq \frac{2r^2}{(1-\frac{2}{5})^2\cdot 14r^2} = \frac{25}{72} \end{aligned} \tag{3.13}$$

As $\text{BREV}(D)$ is at least the revenue obtained by setting the grand bundle at the price $\frac{2}{5}\cdot\text{VAL}(D_{\emptyset}^{C})$, and the probability according to the result above is $1 - \dfrac{25}{72} = \dfrac{47}{72}$.

$$\text{BREV}(D) \geq (\frac{2}{5}\cdot \text{VAL}(D_{\emptyset}^{C}))\cdot\frac{47}{72} = \frac{47}{180}\cdot\text{VAL}(D_{\emptyset}^{C}) > \frac{1}{4}\text{VAL}(D_{\emptyset}^{C}) \tag{3.14}$$

□

Combining proposition 1 and 2 with Lemma 6, we can show our main result is true.

4

Extensions and Limitations

The author further extend their results to multiple buyers with all valuations sampled independently. In this case, their method used in the single buyer auction does have its limitation. As there are many theorems and their proofs are less important compared with the previous one, we only try to list their primary conclusions as simple and plain as possible and prove the theorem which shows their limitation.

1. Selling items separately achieves a **logarithmic** approximation to the optimal revenue. That is,

$$\begin{gathered}\text{for any number of } m \text{ buyers and } n \text{ items,} \\ (2+2e^{1/4}+\ln 4)\mathrm{SREV}(D) \geq \mathrm{REV}(D)\end{gathered} \tag{4.1}$$

2. Selling items separately can achieve a **constant-factor** approximation **unless** the welfare of D is sufficiently well concentrated around its expectation. That is,

$$\begin{gathered}\text{for any } c \geq 4\sqrt{2}, \text{ either } (c+5)\mathrm{SREV}(D) \geq \mathrm{REV}(D) \\ \text{or the welfare of } D \text{ is } (\frac{3}{4}-\frac{24}{c^2}) - \text{concentrated.}\end{gathered} \tag{4.2}$$

3. Bundling together does **not** work when the welfare is highly concentrated, even in the case of $\mathrm{PREV}(D)$, a combination of bundling together and selling separately.

Next, we will give a proof of the theorem shown in equation (4.2). Before starting the proof, we will firstly clarify the definition of *d-concentrated* in the theorem.

Definition. *We say that a one-dimensional distribution F is d-concentrated if there exists a value C such that probability $Pr_{x\sim F}[|x - C| \leq C/2] \geq d$*

Then, we will prove equation (4.2) using the trick similar to section 3.2.

Proof: We will first prove the bound on the revenue from the tail of the distribution D, using a similar choice of t_i. Then, we will establish bound between SREV and REV, subject to the welfare of D not being too concentrated around its expectation.

(1) bounding the revenue from the tail

> **Proposition 3.** For any number of buyers, if $t_i = 4r/r_i$ for all i, then $\sum_A p_A \mathrm{REV}(D_A^T) \leq 5\mathrm{SREV}(D)$

To rewrite $\sum_A p_A \mathrm{REV}(D_A^T)$, we may use equation (4.1). Noting that $6 + \ln n \leq 4n$ for all $n \leq 2$ and using the reordering skill similar to 3.2.1, we can infer that

$$\sum_A p_A \mathrm{REV}(D_A^T) \leq 4 \cdot n \cdot \mathrm{SREV}(D) \leq \sum_A 4p_A|A| \sum_{i\in A} r_i/p_i = 4\sum_i r_i \sum_{A\ni i} |A|p_A/p_i$$

Again, as the value $\sum_{A\ni i} |A|p_A/p_i$ is the expected number of items in the tail, conditioned on i being in the tail. Therefore, using Lemma 3, we can get

$$\sum_{A\ni i} |A|p_A/p_i \leq 1 + \sum_{j\neq i} p_j \leq 1 + 1/4$$

Thus,

$$\sum_A p_A \mathrm{REV}(D_A^T) \leq 4\sum_i r_i \sum_{A\ni i} |A|p_A/p_i \leq 4 \cdot (1 + 1/4) \sum_i r_i = 5\sum_i r_i = 5\mathrm{SREV}(D)$$

Set all $t_i = 4r/r_i$. Combing proposition 3 with Lemma 6 yields

$$5\mathrm{SREV}(D) + \mathrm{VAL}(D_\emptyset^C) \geq \mathrm{REV}(D) \tag{4.3}$$

(2) bounding the welfare from the core

Again, we will use the variance and expectation of $D_\emptyset^C$, divide the case by its expectation and prove the theorem case by case.

case 1: $c \cdot \mathrm{SREV}(D) \geq \mathrm{VAL}(D_\emptyset^C)$, when the bound of theorem holds.

In this case, we have $(c+5)\mathrm{SREV}(D) \geq \mathrm{REV}(D)$ and the claim follows.

case 2: $c \cdot \mathrm{SREV}(D) \leq \mathrm{VAL}(D^C_\emptyset)$, when the welfare of D is $(\frac{3}{4} - \frac{24}{c^2}) -$ concentrated.

According to 3.11, we can recall that $var(D^C_i) \leq 2t_i r_i^2$. Again summing over all i, we get

$$var(D^C_\emptyset) \leq 2\sum_i t_i r_i^2 = 2\sum_i (4r) r_i = 8r^2$$

Thus, we have $var(D^C_\emptyset) \leq 8r^2$ and $\mathrm{VAL}(D^C_\emptyset) \geq cr$. Recalling Chebyshev's inequality, we get

$$Pr[|\sum_i v_i^* - \mathrm{VAL}(D^C_\emptyset)| \geq \mathrm{VAL}(D^C_\emptyset)/2] \leq \frac{var(D^C_\emptyset)}{\mathrm{VAL}(D^C_\emptyset)^2/4} \leq \frac{8r^2}{c^2r^2/4} \leq \frac{32r^2}{c^2r^2} = \frac{32}{c^2}$$

According to definition 1, the welfare of $\mathrm{VAL}(D^C_\emptyset)$ is $(1 - \frac{32}{c^2}) -$ concentrated.

To show the welfare of D is $(\frac{3}{4} - \frac{24}{c^2}) -$ concentrated, we can observe that $\vec{v}$ is sampled in the support of $D^C_\emptyset$ with probability exactly $\prod_i (1 - p_i)$. Based on Lemma 3, we know

$$\sum_i p_i \leq \sum_i \frac{1}{t_i} \leq \sum_i \frac{r_i}{4r} = \frac{\sum_i r_i}{4\sum_i r_i} = \frac{1}{4}$$

As each $p_i \geq 0$, this is minimized when exactly one p_i is $1/4$ and all the others are 0. Thus, $\prod_i (1 - p_i) = 3/4$. That is, $\vec{v}$ is in the support of $D^C_\emptyset$ with probability at least $3/4$. Hence, the welfare of D is $\frac{3}{4} \cdot (1 - \frac{32}{c^2}) -$ concentrated, namely $(\frac{3}{4} - \frac{24}{c^2}) -$ concentrated.

□

Hence, the biggest limitation to extend the result to multiple buyers is to **tackle the concentration of welfare to get a constant-factor approximation**. The intuition is that the allocation of items to buyers which provides such welfare can change dramatically between realizations. The subsequent work of Yao[5] has succeeded in doing that with the VCG mechanism (which sells each item separately using a second-price auction), but maintain the entry fee (from another perspective, bundling together can be seen as the mechanism which gives the buyer all items for free if the buyer pays an entry fee to participate).

Partition mechanisms $\mathrm{PREV}(D)$ can be defined as follows: the optimal revenue obtained by any partition mechanism when the buyer profile is drawn from D. That is, the maximal revenue obtained by first partitioning the items into disjoint bundles, and then running Myerson's optimal auction separately for each bundle, treating each bundle as a single item. To better understand it, we explore the connection between SREV(D), BREV(D), and PREV(D) for multiple buyers. Their relationships are as follows:

- $\max\{\text{SREV}(D), \text{BREV(D)}\}$ achieves a constant-factor approximation to PREV(D) when **either** buyers or items are i.i.d.
- $\max\{\text{SREV}(D), \text{BREV(D)}\}$ achieves at best an $\Omega(\ln(n))$-approximation to PREV(D) when **neither** buyers nor items are i.i.d.

For one buyer with correlated results, prior work of [6, 7] has shown that there is no hope of obtaining a non-zero bound between any of SREV(D), $\max\{\text{SREV}(D), \text{BREV}(D)\}$ and PREV(D) and REV(D). Yet it is still important to understand the relationship between theses mechanisms of varying complexity.

- For any correlated distribution D for a single buyer and n items, SREV(D) is a $O(\log n)$ approximation to BREV(D), and thus also to REV(D).
- The bound above is tight, taking $\max\{\text{SREV}(D), \text{BREV}(D)\}$ can't guarantee anything better.

5 Acknowledgements and Harvest

I would like to express my gratitude to my teacher, Zhengyang Liu, who teaches with great care and warmth and is always willing to improve his teaching skills. Through this course, I have a preliminary understanding of Algorithmic Game Theory and also gain a lot in the process of finishing the report.

The initial motivation for me to choose this paper is that it is highly-cited and published relatively not long ago compared with others, which are symbols of a paper worth reading. After the first pass of it, I find it is an extension of the revenue-maximizing auctions we have learned in class. As the single item auction is only an ideal scenario in real auctions, extending it to multi-item case indeed has its practical value. Then, I start digging into the article. In the process, I am amazed to find that I am able to read a long academic paper with many math theorems as long as I put efforts into it and use proper skills (I used to read a paper from beginning to end with much wasted effort!). I grasp the meaning of the whole proof by its elegant derivation. Not only that, I learn to understand these theorems by intuition and am impressed by the combination of the theorem itself and its practical meaning.

I have tried my best to show my sincerity in both my homework and report: I make the proof as rigorous as possible in my homework and narrate the paper as plain and logical as possible using my own understanding. I would be more than grateful if you are so kind to grade me 100 (if not, that's also ok). Thanks again!

References

[1] Moshe Babaioff, Nicole Immorlica, Brendan Lucier and S Matthew Weinberg. 'A simple and approximately optimal mechanism for an additive buyer'. In: *Journal of the ACM (JACM)* 67.4 (2020), pp. 1–40 (cit. on p. ii).

[2] Sergiu Hart and Noam Nisan. 'Approximate revenue maximization with multiple items'. In: *Journal of Economic Theory* 172 (2017), pp. 313–347 (cit. on pp. 2, 9).

[3] Xinye Li and Andrew Chi-Chih Yao. 'On revenue maximization for selling multiple independently distributed items'. In: *Proceedings of the National Academy of Sciences* 110.28 (2013), pp. 11232–11237 (cit. on pp. 6, 9).

[4] Yang Cai and Zhiyi Huang. 'Simple and nearly optimal multi-item auctions'. In: *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM. 2013, pp. 564–577 (cit. on p. 9).

[5] Andrew Chi-Chih Yao. 'An n-to-1 bidder reduction for multi-item auctions and its applications'. In: *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*. SIAM. 2014, pp. 92–109 (cit. on p. 16).

[6] Patrick Briest, Shuchi Chawla, Robert Kleinberg and S Matthew Weinberg. 'Pricing lotteries'. In: *Journal of Economic Theory* 156 (2015), pp. 144–174 (cit. on p. 17).

[7] Sergiu Hart, Noam Nisan et al. *The menu-size complexity of auctions*. Center for the Study of Rationality, 2013 (cit. on p. 17).