SINGLE BIDDER MULTI-ITEM AUCTIONS

FRANKLYN WANG AND ERIC ZHANG

Abstract. Here, we describe the problem of maximizing revenue in single-bidder auctions, where the valuation of a bidder is drawn randomly from some known distribution. This is a well-studied problem, beginning with [9], which studied auctions with a single item. In that case, the optimal auction is straightforward; the seller posts a single take-it-or-leave-it price. However, the problem gets much more complicated when considering the case of multi-item auctions (even for two items with independent value distributions). For a long time, there was little progress in finding the optimal mechanism, until [3], which used LP duality to show that optimal mechanisms can be very complicated in some cases. At the same time, a large branch of economics emerged that was focused on achieving *simplicity* in mechanism design. We look at a few of these approaches, most notably those in [5], [2], and [6].

1. Introduction

In mechanism design, an important problem (and arguably the most important problem) is to maximize revenue for the seller of a set of objects. In this paper, we focus on mechanisms with a single buyer. A result by Myerson ([9]) shows that in the case of a single-item auction with a buyer whose valuation is drawn uniformly from $[0, 1]$ (but whose exact value is unknown to the seller), the optimal auction mechanism that maximizes the seller's profits is a take-it-or-leave-it offer at a price of $1/2$.

This is the simplest case, with exactly one good. Logically, it must also be true that if there are two goods, and a buyer's valuation is drawn i.i.d. from $[0, 1]$ for both goods, then the optimal mechanism should be to sell each of them independently at a price of $1/2$, right? **Wrong!**

While the revenue-maximizing mechanism is quite complicated, consider for example the mechanism that sells both goods combined at a price of $3/4$. The probability that the valuation is above $3/4$ is $23/32$. This leads to an expected revenue of $69/128$, which is larger than $1/2$.

The fact that selling the items together achieves a better result than separately, even when the items' prices are independent, is significant. Finding the optimal mechanism's revenue (even for this case) turns out to be extremely challenging. The optimal mechanism (from Gonczarowski [5]) is to provide two options:

(1) Buy any single item for a price of $2/3$.
(2) Buy both items for the combined price of $(4 - \sqrt{2})/3$.

This mechanism is substantially more complicated than one may have expected. Although it does not involve selling probabilistic lottery tickets for items, or infinite menus, it still involves a selection from multiple options, as well as bundling several items together. This suggests that even in the simplest case of i.i.d. uniform buyer price distributions, there are very complicated optimal mechanisms.

In fact, [4] shows that for some two-item auctions, the revenue-maximizing mechanism requires using infinitely many choices, and some optimal constructions require probabilistic choices (e.g., you can bid 0.4 for a 0.7 chance at receiving an item).

Being perfect is hard, and in cases where infinitely many choices are needed, is essentially impossible. So what can we do? A recently developing field of economics focuses on simplicity, a parameter whose need is becoming readily apparent in an ever more complex world. Myerson's auction is very simple. What are some other simple mechanisms we can try? We can try selling each item separately, or sell all the items together in a bundle. [2] shows that one of these two mechanisms is always guaranteed to generate at least $1/6$ the optimal revenue. However, in practical applications, even this result about optimality up to a constant factor is slightly unsatisfying. What merchant was ever satisfied with $1/6$ of the optimal revenue?

To this end, [3, 4] show ways of obtaining true optimality via a linear programming approach. [5] uses this duality to prove lower bounds on the complexity of a mechanism that is needed to obtain the optimal revenue minus ϵ. Another paper, [1] proves an upper bound on the complexity of a mechanism that is needed to obtain all but ϵ of the maximum revenue by an explicit construction; these bounds are sufficient to prove an exact bound on the communication complexity of any mechanism which obtains all but ϵ of the maximum revenue.

2. Model

For any simply connected subset S of $\mathbb{R}^n$, we define $\Delta(S)$ to be the set of all probability distributions supported on S.

We describe a model of the single-bidder mechanism design problem. An auctioneer has a selection of n goods that he wants to sell, and he wants to design a mechanism that maximizes his revenue (as the goods have no cost to him). He has a single bidder, whose valuations for the items are drawn from a distribution $\mathcal{D} \in \Delta([0,1]^n)$. The auctioneer's goal is to maximize his expected revenue from selling the goods, given control over the pricing mechanism. The bidder's objectives are to maximize his utility, which is linear in his value of the items (i.e., he is risk neutral; a lottery ticket that has value v with probability $1/2$ and value 0 with probability $1/2$ has overall value $v/2$). In effect, the bidder's strategy places constraints on the auctioneer's mechanism design.

The mechanism plays out as follows. First, the bidder is given confidential knowledge of a vector $z \in [0,1]^n \sim \mathcal{D}$ of his valuations for each item, drawn according to a public probability distribution. He reports some valuations z' to the auctioneer, who in turn gives the bidder a bundle: consisting of certain probabilities $x(z') \in [0,1]^n$ of obtaining each item, as well as price $p(z') \in \mathbb{R}$ that is charged in total. Since the bidder's utilities are additive over the objects and linear, the bidder's utility is given by

$$u(z, z') = x(z') \cdot z - p(z'),$$

where $x(z')$ is the bundle (a vector of probabilities denoting the probability of receiving the object), z is the valuation (a vector containing the values of each of the objects), and $p(z')$ is the price that the bidder pays.

We restrict ourselves only to mechanisms that are *incentive compatible* and *individually rational*, meaning that the bidder will always be incentivized to report his true valuations for each item, and he will always gain positive expected utility from participating in the mechanism. This means that for all z and z',

$$x(z) \cdot z - p(z) \geq \max\{x(z') \cdot z - p(z'), 0\}.$$

By the revelation principle in mechanism design, these incentive-compatible direct mechanisms are suitable to maximize revenue. Consequently, all of these mechanisms boil down to a single kind: the auctioneer describes a menu of bundles, which each include n probabilities of receiving the individual items, as well as the overall price of the bundle. The bidder selects the bundle which gives them the most expected utility.

However, in the rest of this chapter we will frequently find it useful to move between these two descriptions of the problem: one in terms of discrete menus of bundles, and one in terms of an allocation function x (hence *allocation rule*) and a price function p (hence *price rule*). In general, it will be easier to do calculations with the menu interpretation, and simpler to express the duality framework with the utility function $u(z) = x(z) \cdot z - p(z)$.

In what follows, suppose that we are given a (possibly multidimensional) distribution with cumulative distribution function F and density function f, supported on $\mathbb{R}^n$. Many times we will consider the case where the marginal distributions $F_1, F_2, \ldots F_n$ for each item are independent; we refer to this as the *independent setting*. Let $\mathrm{OPT}(F)$ be the maximum revenue that one can obtain in a single-bidder auction given a buyer whose value distribution follows F. Also, let $\mathrm{REV}(F, M)$ be the revenue obtained by mechanism $M = (x, p)$ on the distribution F.

3. Simplicity

Before we attempt to achieve all but ϵ of the optimal revenue, we can explore how close we may get using very simple mechanisms. If simple mechanisms can yield revenues exceeding $\mathrm{OPT}(F)-\epsilon$, we would not require more complicated menus. In this section, we focus on the independent setting.

Two broad types of simple mechanisms are considered in [2]:

(1) *Grand-bundling mechanisms*, where the a single bundle of all the goods is offered together. Denote the optimal revenue from this case as $\mathrm{OPTG}(F)$.
(2) *Posted-price mechanisms*, where each item is assigned a price, and buyer is allowed to buy each separately. Denote the optimal revenue from this case as $\mathrm{OPTP}(F)$.

These mechanisms are both simple enough to be easily analyzed. One result is that both of these can be arbitrarily bad compared to the optimal mechanism. To illustrate this, we show, in the following two lemmas, that both OPTG and OPTP can be arbitrarily bad compared to each other.

Lemma 3.1 ([6])**.** *For any $\epsilon > 0$, there exists F in the independent setting so that*

$$\frac{\mathrm{OPTG}(F)}{\mathrm{OPT}(F)} \leq \frac{\mathrm{OPTG}(F)}{\mathrm{OPTP}(F)} < \epsilon.$$

Proof of Lemma 3.1. Consider n goods, so that good i has value 2^i with independent probabilities 2^{-i} from the buyer and zero probability otherwise.

The posted price mechanism here sells each good for 2^i, which gives an expected revenue of n. This is because selling for any lower price would lead to lower value with the same probability, while selling for a higher price would make the mechanism no longer individually rational. On the other hand, to analyze the grand bundle mechanism (selling all the goods together), suppose that they are sold at a price p. Observe if that value $\|z\|_1$ of the bundle is worth at least p, there must be a good of value at least $p/2$. By a union bound, the probability that the value of the bundle is worth more than p is at most

$$2^{-\lfloor \log_2(p) \rfloor + 1} + 2^{-\lfloor \log_2(p) \rfloor} + 2^{-\lfloor \log_2(p) \rfloor - 1} + \cdots < \frac{4}{p},$$

implying a maximum revenue of 4. Therefore, by increasing the number n of goods, we can construct distributions F such that

$$\frac{\mathrm{OPTG}(F)}{\mathrm{OPTP}(F)}$$

is arbitrarily small. □

Lemma 3.2 ([6])**.** *For any $\epsilon > 0$, there exists a valuation distribution F so that*

$$\frac{\mathrm{OPTP}(F)}{\mathrm{OPT}(F)} \leq \frac{\mathrm{OPTP}(F)}{\mathrm{OPTG}(F)} < \epsilon,$$

even if we restrict F to be the product distribution of i.i.d. random variables.

Proof of Lemma 3.2. Note that a distribution with finite expected value will be difficult to make work here, because $\mathrm{OPTG}(F)$ will tend to concentrate around the expected value of the sum $\|z\|_1$ by the law of large numbers. Instead, we consider the distribution which might be the worst possible for the posted price mechanism: the equal revenue distribution, where any posted price yields the same revenue for that item. In this case, while the grand bundle benefits from independent concentration, the posted-price mechanism does not. Formally, consider the distribution on the price of the first good, z_0, with cumulative distribution function F given by

$$F(z_0) = \begin{cases} 1 - 1/z_0 & z_0 \geq 1 \\ 0 & z_0 < 1 \end{cases}$$

Observe that this distribution has a very thick tail. For a single item in this distribution, if it is sold any price $p \geq 1$, then the expected revenue is given by

$$p \cdot \Pr(z_0 > p) = p \cdot (1 - F(p)) = 1.$$

Thus, the expected revenue from the optimal posted-price mechanism on this distribution is n given n independent and identically distributed items at this valuation.

Now, given i.i.d. valuations $X_1, \ldots, X_n \sim F$, we want to show that $S_n := X_1 + X_2 + \ldots + X_n$ achieves high values (in particular, $\omega(n)$) with constant $\Theta(1)$ probability. This would allow a grand bundle mechanism priced at that value to achieve a better asymptotic expected revenue as n increases. However, this will be hard to do with central limit theorem or law of large numbers, since the expected value of the distribution of X_i is

$$\int_1^\infty x f(x)\, dx = \int_1^\infty x \left(\frac{1}{x^2}\right) dx = \infty.$$

Instead, we will use Chebyshev's inequality to show that S_n is $\Omega(n \ln n)$ with constant probability. Since the variances of X_i are also not defined, we clip off the tails at some value $M > 1$ and take instead $X_1' = \min(X_1, M)$. Then, the expected value of X_1' is

$$\mathbb{E}[X_1'] = \int_1^M x f(x)\, dx + M \cdot \frac{1}{M} = \ln M + 1,$$

and the variance is

$$\mathrm{Var}[X_1'] = 2M.$$

By the additivity of variances of independent random variables, the sum of X_i' has a mean of $n \ln M + n$ and a standard deviation of $\sqrt{2Mn}$. Letting $M = n$, note that this distribution will with nonvanishing probability have a sum greater than $n \ln n - 4n$ by Chebyshev's inequality. Specifically,

$$\Pr(S_n < n \ln n - 4n) < \Pr(S_n' < n \ln n - 4n) < \Pr(|S_n' - (n \ln n + n)| < 5n) = \frac{2n^2}{25n^2} = \frac{2}{25},$$

and thus the expected revenue is bounded below by $\frac{23}{25}(n \ln n - 4n) = \Omega(n \ln n)$. □

Reflecting on what makes these mechanisms arbitrarily suboptimal, the posted price mechanism loses to the grand bundling mechanism because when n is large and the variables are i.i.d., as the sum of all the items as a whole tends to be concentrated around the expected value by the law of large numbers. Intuitively, the grand bundling mechanism helps reduce risk for the seller. This trend even holds when the distributions, such as above, have infinite expected value and undefined variance (and in fact, only gets stronger).

Now, here's a very surprising theorem that shows that one of our two simple mechanisms is always within a constant factor of the optimal mechanism:

Theorem 3.3 ([2]). *We have that*

$$6 \cdot \max(\mathrm{OPTP}(F), \mathrm{OPTG}(F)) \geq \mathrm{OPT}(F)$$

This is a surprising result because it's quite counterintuitive that while either mechanism can give arbitrarily bad approximations, the better of the two will always be within a constant factor of the optimal.

Proving this theorem is difficult, but the main idea is to break the distribution into a "core" and a tail, where the core contains almost all of the probability mass, and the tail contains the points of large revenue. We will show a specific result about this, which we call the core decomposition lemma.

However, before we can prove the core decomposition lemma, we need a few intermediate results. Consider arbitrary distributions F, F' each supported on $\mathbb{R}^k$ for some dimensions k, and define $\mathrm{VAL}(F)$ to be the expected valuation of the buyer on the entire distribution. We have the following bound on the optimal revenue of a mechanism.

Lemma 3.4 ([6])**.** *If F is independent from F', then we have that*

$$\mathrm{OPT}(F \times F') \leq \mathrm{VAL}(F) + \mathrm{OPT}(F').$$

Proof. Consider a mechanism M which achieves the optimal revenue on $F \times F'$. We will use this to construct a mechanism on F' that has expected revenue at least $\mathrm{OPT}(F \times F') - \mathrm{VAL}(F)$, which will imply the lemma.

Consider M', the restriction of M to elements of F', where a mechanism function (x, p) is restricted to

$$x(z) \longmapsto x(z)|_{F'},$$
$$p(z) \longmapsto p(z) - z \cdot x(z)|_F.$$

In particular, for each element f_i in F that is included in the bundle with probability x_i, we discount the price of the restricted bundle in the new mechanism M' by the sum the expected valuation $x_i z_i$. This mechanism retains the IC and IR conditions, and on expectation decreases the revenues of the auctioneer by

$$\mathbb{E}[z \cdot x(z)|_F] \leq \mathrm{VAL}(F),$$

which yields the result. □

The key idea in the above was discounting the bidder $\sum_i x_i z_i$ in money. We could also have tried discounting the the bidder by exactly $\sum_i z_i$ in money, but that might have broken incentive compatibility.

We now prove another lemma, which suggests that breaking up a distribution into many pieces makes it easier to optimize revenues. Intuitively, the reason this is true is that breaking up a distribution into pieces and optimizing revenues over those distributions allows you to specialize to each piece.

Lemma 3.5. *Given some finite set A, assume that for all $a \in A$, we have $p_a \in \mathbb{R}$ and $f_a : D \to \mathbb{R}$, where D is the support of F, such that*

$$f(x) = \sum_{a \in A} p_a f_a(x).$$

In other words, f is a mixture of distributions f_a. Then,

$$\mathrm{OPT}(F) \leq \sum_{a \in A} p_a \, \mathrm{OPT}(F_a).$$

Proof. Let M be any mechanism on F. Note that

$$\mathrm{REV}(F_a, M) < \mathrm{OPT}(F_a),$$

so we have

$$\sum_{a\in A} p_a \operatorname{REV}(F_a, M) \leq \sum_{a\in A} p_a \operatorname{OPT}(F_a).$$

This implies that

$$\operatorname{REV}(F, M) \leq \sum_{a\in A} p_a \operatorname{OPT}(F_a),$$

as desired. □

Now, our next goal is to prove a core decomposition lemma that is a key step in proving Theorem 3.3. For each of the items $i = 1, 2, \ldots n$, let r_i be the revenue attained by an optimal single-item posted-price auction for that individual good. Let $t_i \geq 1$ be a set of n parameters, and define the *core* of the distribution to be the product of the n intervals

$$F_\emptyset^C = \prod_{i=1}^{n} [0, t_i r_i].$$

Intuitively, the idea behind this definition is that for *any* mechanism, the probability that a single item i "contributes" (in a vague sense) more than $t_i r_i$ to the revenue is at most $1/t_i$. We can now introduce some notation to formalize the core decomposition idea that we stated earlier:

- Let p_i be the probability that the bidder's valuation z_i for item i is greater than $t_i r_i$, the endpoint of the corresponding interval in the core.
- Let F_i^C be the core of item i, or the conditional distribution of the buyer's valuation z_i given that $z_i \leq t_i r_i$.
- Let F_i^T be the tail of F_i, which is the conditional distribution of z_i conditioned on $z_i > t_i r_i$.
- A is an arbitrary subset of the n items being sold, which we usually take to be the set of items whose valuations for the bidder are in the tails.
- F_A^T: A is a subset of items, and F_A^T is a product distribution equal to $\prod_{i\in A} F_i^T$
- F_A^C: A is a subset of items, and F_A^C is a product distribution equal to $\prod_{i\notin A} F_i^C$.
- F_A: $F_A^C \times F_A^T$.
- p_A represents the probability that the bidder's valuation is in the support of F_A.

With this notation, we are now ready to prove the core decomposition lemma (the main lemma in [2]), which is the key step in proving Theorem 3.3.

Lemma 3.6 (Core Decomposition Lemma)**.**

$$\operatorname{OPT}(F) \leq \operatorname{VAL}(F_\emptyset^C) + \sum_{A\subseteq\{1,\ldots,n\}} p_A \operatorname{OPT}(F_A^T).$$

Proof of Lemma 3.6. In light of the above, this quickly follows. We know that

$$\operatorname{OPT}(F) \leq \sum_A p_A \operatorname{OPT}(F_A) \leq \sum_A p_A(\operatorname{OPT}(F_A^T) + \operatorname{VAL}(F_A^C)),$$

and the result follows by using the bound

$$\operatorname{VAL}(F_A^C) \leq \operatorname{VAL}(F_\emptyset^C).$$

□

This concludes the setup and proof of the core decomposition lemma. For brevity, we do not present a complete proof of Theorem 3.3; this can be found in [2]. However, we will provide a rough sketch of the argument.

To go from Lemma 3.6 to Theorem 3.3, we need to choose the values of t_i correctly, after which these terms in the core and tails will be reasonable to bound. If the values of t_i are all relatively large, there generally will not be any values in the tail, which will allow us to bound them. The final technical step is to show that a grand bundling approach provides most of the revenue in the core, while a single item approach provides most of the revenue in the tails.

Note that one difficulty this approach faces is that if the values of t_i are all the same, then the core becomes highly skewed. Thus, we end up choosing $t_i = cr/r_i$, where r is the maximum revenue obtainable by a grand bundling mechanism. This makes $t_i r_i = cr$ for all i, so the core is a cube.

This concludes our discussion, setup, and sketch of the proof for Theorem 3.3. A natural next question would be if more sophisticated analyses could prove a tighter bound on the efficacy of this simple mechanism. However, it turns out that this strategy can only go so far, as the following lemma shows.

Lemma 3.7 ([2]). *There exists a distribution F for which*

$$2 \cdot \max(\mathrm{OPTP}(F), \mathrm{OPTG}(F)) \leq \mathrm{OPT}(F).$$

Thus, this simple mechanism, while producing asymptotically optimal results, does not provide an arbitrarily close approximation scheme.

4. Duality

In this section, we attempt to find, in a general setting, the mechanism that extracts the maximum value.

Recall from the revelation principle that any such mechanism can be described as follows. The bidder announces his values z. Then a mechanism $M = (x, p)$ consists of an an allocation rule $x(z)$ providing probabilities of receiving each item in the bundle, as well as a price rule $p(z)$ that represents the posted price for the bundle. The mechanism M provides the buyer with $u(z) = x(z) \cdot z - p(z)$ in utility.

We wish to maximize the expected revenue obtained by the seller, which is the integral

$$\int p(z) f(z)\, dz$$

of the price rule. However, we cannot simply choose any price rule in this optimization problem. For an extreme example, the price rule $p(z) = 100$ and $x(z) = \mathbf{0}$ is not valid. There is simply no reason for the buyer to participate in this mechanism, where he pays 100 units of money for a bundle of zero goods. In the language of game theory, this mechanism is not *individually rational*. We have to find a way to encode individual rationality and incentive compatibility in the language of the price rule and allocation rule.

In terms of u, these expressions become slightly more palatable. Incentive compatibility is true if and only if u is non-decreasing, convex, continuous, and $\nabla u(z) \in [0, 1]^n$ (in other words, 1-Lipschitz), a result shown in [7].

We can now give the expected revenue as an integral that we wish to maximize, expanding it using the expression for u as

$$\int p(z)f(z)\,dz = \int (x(z)\cdot z - u(z))f(z)\,dz.$$

We can now calculate that $x(z) = \nabla u(z)$ by the envelope theorem, so we want to maximize

$$\int (\nabla u(z)\cdot z - u(z))f(z)\,dz.$$

By standard multivariable calculus (c.f. [3]), we can rewrite this as

$$\int u(z)(-\nabla f(z)\cdot z - (n+1)f(z))\,dz.$$

Thus, we want to take a function $u(z)$ that maximizes this value. Here, let $d_0 = \mathbf{0}$ be the origin. Then, we can split $D = \mathrm{supp}(F)$ into regions of positive and negative contribution to this integral by defining two sets:

$$\mathcal{X} = \{z \in D \mid -\nabla f(z)\cdot z > (n+1)f(z)\} \setminus \{d_0\},$$
$$\mathcal{Y} = \{z \in D \mid -\nabla f(z)\cdot z \le (n+1)f(z)\} \setminus \{d_0\}.$$

Using the expressions above, we can define density functions $\mu_d^{\mathcal{X}}(x) = -\nabla f(x)\cdot x - (n+1)f(x)$ and $\nu_d^{\mathcal{Y}}(y) = (n+1)f(z) + \nabla f(y)\cdot y$. By the Lebesgue-Radon-Nikodym theorem, we can define unique measures $\mu^{\mathcal{X}}$ and $\nu^{\mathcal{Y}}$ with these densities. Therefore, we want to find a function which maximizes

$$\int_{\mathcal{X}} u(x)\mu_d^{\mathcal{X}}(x)\,dx - \int_{\mathcal{Y}} u(y)\nu_d^{\mathcal{Y}}(y)\,dy.$$

At this point, this problem becomes an optimization problem, where we want to find a utility $u(z)$ that maximizes the objective. $\mathcal{X}$ and $\mathcal{Y}$ can both be calculated purely from the density function.

We will actually slightly "patch" $\mathcal{X}$ by extending it with the origin d_0 as an extra element, such that $\mu^{\mathcal{X}_0}(\{d_0\}) = 1$. Because $\mu^{\mathcal{X}}(\mathcal{X}) - \nu^{\mathcal{Y}}(\mathcal{Y}) = -1$ (c.f. [3]), working with $\mathcal{X}_0$ and $\mathcal{Y}$ will give us the nice property that $\mu^{\mathcal{X}_0}(\mathcal{X}_0) = \mu^{\mathcal{Y}}(\mathcal{Y})$.

However, it will in general be very difficult to solve this type of problem. Instead of attempting to optimize given the convexity and 1-Lipschitz constraints on u, which are natural but hard to optimize over, we use a different type of constraint. Define a "cost function" $c : \mathcal{X}_0 \times \mathcal{Y} \to \mathbb{R}$ by summing over the differences of corresponding vector coordinates,

$$c(x,y) := \sum_{i=1}^{n} \max\{x_i - y_i, 0\}.$$

Note by the 1-Lipschitz condition and the monotonicity mentioned earlier, $u(x) - u(y) \le c(x,y)$. If we optimize the objective under this weaker constraint, and the optimizing function also happens to satisfy the original constraints of incentive compatibility and individual rationality, we are done. In fact, [4] shows that optimal solutions for this cost constrain essentially always satisfy the original constraints.

Now, to bound the objective function subject to cost constraints, we use a coupling of finite measures. Let $\Gamma(\mu^{\mathcal{X}_0}, \nu^{\mathcal{Y}})$ be the set of joint distributions $\gamma(x,y)$ with marginals equal to $\mu^{\mathcal{X}_0}$ and $\nu^{\mathcal{Y}}$. Then, a coupling argument tells us that the objective function satisfies

$$\mathrm{OPT}(F) = \int_{\mathcal{X}_0 \times \mathcal{Y}} u(x) - u(y)\, d\gamma(x,y) \leq \inf_{\gamma \in \Gamma(\mu^{\mathcal{X}_0}, \nu^{\mathcal{Y}})} \int_{\mathcal{X}_0 \times \mathcal{Y}} c(x,y)\, d\gamma(x,y).$$

This is exactly an optimal transport problem [8], which here is expressed as the dual of our optimization problem. This is how we can equate, via weak duality, the revenue maximization problem to the optimal transport problem.

5. Menu Sizes

In this section, we move on to an analysis of the simplicity of arbitrarily close approximation procedures for optimal mechanisms. Using the duality framework described in the past section, we construct a particular example of an auction that is hard to approximate with using a finite menu size. This will provide a general lower bound on the simplicity (menu size) of approximation algorithms of a given error threshold. The main result is stated below.

Theorem 5.1 ([5], Theorem 2)**.** *There exists $C(\epsilon) = \Omega\left(1/\sqrt[4]{\epsilon}\right)$ and a distribution $F \in \Delta([0,1])$, such that for every $\epsilon > 0$ it is the case that* $\mathrm{REV}(F \times F, M) < \mathrm{OPT}(F \times F) - \epsilon$ *for every mechanism with menu-size at most $C(\epsilon)$.*

In other words, if we want to come within an ϵ of the optimal revenue, we need at least $\Omega(1/\sqrt[4]{\epsilon})$ items on the menu. Before we prove why $\Omega(1/\sqrt[4]{\epsilon})$ items are needed, let's ask the question: why does this number scale exponentially with $\log(1/\epsilon)$? A reasonable explanation is that with any finite menu, the decision boundaries that the buyer makes will appear to be piecewise linear. On the other hand, if the optimal decision boundaries are very "round", it may be very difficult to approximate it well with lines. This forms the intuition for the proof of the complexity lower bound.

A key step in proving this result will be the notion of *strong duality*. Consider some problems p and d, which are primal and dual. Let o_p and o_d be real-valued functions defined across S_p and S_d respectively. The primal is a question of *maximizing* a function o_p over a set S_p, and the dual is a question of *minimizing* a function o_d over a set S_d. Then, a weak duality relation states that for all $s_p \in S_p$ and $s_d \in S_d$,

$$o_p(s_p) \leq o_d(s_d).$$

For example, we showed in the last section that the mechanism design (p) and optimal transport (d) problems are weakly dual, subject to appropriate constraints.

A *strong duality* theorem, on the other hand, states that in addition to a weak duality relation, there always exists a pair $(\hat{s}_p, \hat{s}_d)$ so that

$$o_p(\hat{s}_p) = o_d(\hat{s}_d).$$

In other words, the *duality gap* $\inf_{s_d} o_d(s_d) - \sup_{s_p} o_p(s_p)$ is zero. Importantly, note that finding a pair $(\hat{s}_p, \hat{s}_d)$ immediately solves both the primal and the dual. The result of [4] tells us that not only is the revenue maximization problem weakly dual to a problem in optimal transport, it is also strongly dual.

What we want to show here, then, is that for a certain distribution (in this case, $F \sim \mathrm{Beta}(1,2)$), thinking about the dual will allow us to show that with a limited number of menu items, we are *very* far away from finding the optimal solution (in particular, we come ϵ short of maximizing the expression). As we've shown above, for any coupling $\gamma \in \Gamma(\mu^{\mathcal{X}_0}, \nu^{\mathcal{Y}})$

(especially the one that minimizes the right-hand side), we have

$$\int u\,d\mu^{\mathcal{X}_0} - \int u\,d\nu^{\mathcal{Y}} \leq \int_{[0,1]^2\times[0,1]^2} (u(x)-u(y))\,d\gamma(\mu^{\mathcal{X}_0},\nu^{\mathcal{Y}})$$
$$\leq \int_{[0,1]^2\times[0,1]^2} c(x,y)\,d\gamma(\mu^{\mathcal{X}_0},\nu^{\mathcal{Y}}).$$

For the strong duality, we need *complementary slackness* conditions to hold: in this case, $c(x,y) = u(x) - u(y)$ almost everywhere on γ. What we want to do now is to show that very often, $u(x) - u(y) < c(x,y)$ by a large margin. We can calculate that $d\mu^{\mathcal{X}_0} - d\nu^{\mathcal{Y}}$ is

$$f(x_1)f(x_2)\left(\frac{1}{1-x_1} + \frac{1}{1-x_2} - 5\right)\,dx_1\,dx_2.$$

where $f(x_1)$ and $f(x_2)$ are the density functions of Beta$(1,2)$.

Consider the graph of $\gamma(x,y)$, showing how x and y are coupled. We will not describe the coupling in full detail, but we outline its important properties. In the region given by $\mathcal{R} = [0,x'] \times [0,1]$, the coupling sends points in $\mathcal{A}$ downwards, and in $\mathcal{Z}$ only sends mass from 0 to other points in $\mathcal{Z}$. Some mass here will be transferred outside of region $\mathcal{R}$.

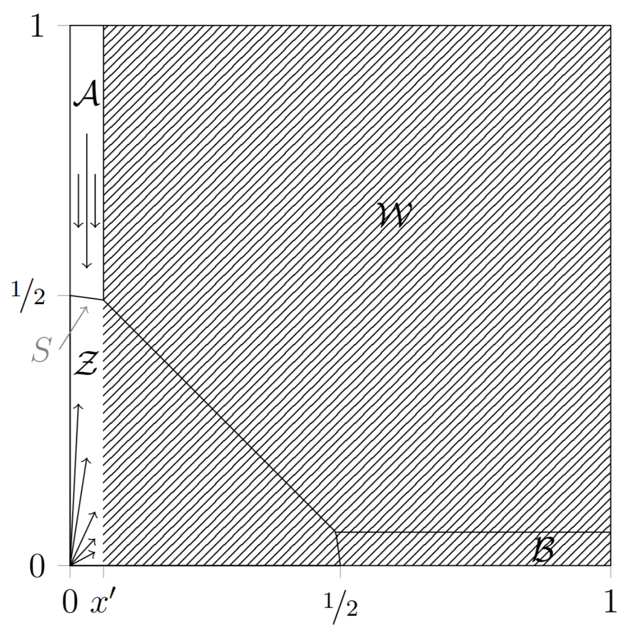

FIGURE 1. Diagram of the coupling on $[0,1]^2$, from [4].

The idea in the problem is that $u(x)-u(y)$ will often have a value which is much less than $c(x,y)$. In particular, in the above graph, equality holds exactly because $u(x)-u(y) = c(x,y)$ for all values of x, y that are coupled under $\gamma(x,y)$. In the above, in region $\mathcal{A}$ the second item is always allocated to the bidder, and in region $\mathcal{Z}$ the second item is never allocated to the bidder (in fact, no items are). The boundary curve S between $\mathcal{A}$ and $\mathcal{B}$ satisfies the equation

$$x_2 = \frac{2-3x_1}{4-5x_1}.$$

Note that this boundary is nonlinear.

Now, consider the boundary curve T, which is the analogue of S formed by the finite-length menu mechanism. One potential problem that can emerge here is that the allocation formed by our menu may be probabilistic in nature. Thus, let T be the smallest point for which the probability of allocating item 2 is above 1/2. Thus, define

$$T(x_1) = \inf\{x_2 \in [0,1] \mid u_2'(x_1, x_2) > 1/2\}.$$

If we restrict our attention to a small neighborhood of the curve, we can obtain a region where the densities are all at least $-d$. Then the idea of the problem is now that T must δ away from S at several points. To state this theorem formally requires tools from differential geometry. Specifically, we state the following result (without proof).

Theorem 5.2 ([5], Proposition 5)**.** *Let $S : [0, x'] \to \mathbb{R}$ be a concave function with radius of curvature at most r everywhere, for some $r < \infty$. For small enough δ, the following holds. For any piecewise-linear function $T : [0, x'] \to \mathbb{R}$ composed of at most $x'/8\sqrt{r\delta}$ linear segments, the Lebesgue measure of the set of coordinates $x_!$ with $|S(x_1) - T(x_1)| > \delta$ is at least $x'/2$.*

Now we have the necessary ingredients to prove Theorem 5.1.

Proof of Theorem 5.1. Now, assume that $T(y_1) - S(y_1) > \delta$. Let $y_2 \in (S(y_1), S(y_1) + \delta/2)$ and let x_2 be so that $\hat{\gamma}$ transfers positive mass from (y_1, x_2) to (y_1, y_2) (and moreover, this is the only way for mass to be transferred. By the definition of T, we have

$$u(x) - u(y) \leq x_2 - y_2 - \delta/4 = c(x, y) - \delta/4.$$

Now consider the case in which $S(y_1) - T(y_1) > \delta$. Let $y_2 \in (S(y_1), S(y_1) + \delta/2)$ and let x_2 be so that $\hat{\gamma}$ transfers positive mass from $x = (0, 0)$ to $y = (y_1, y_2)$. (All mass transferred to y by $\hat{\gamma}$ is from the point x.) By the definition of T,

$$u(x) - u(y) \leq -\delta/4 = c(x, y) \leq -\delta/4.$$

By the technical lemma shown earlier, the set of coordinates y_1 with $|S(y_1) - T(y_1)| > \delta$ has Lebesgue measure at least $x'/2$. Therefore, the inequality originally mentioned must be off by at least $\delta/4 \cdot x'/2 \cdot \delta/2 \cdot d$, which is the density of μ (specifically $d\mu^{\mathcal{X}_0} - d\nu^{\mathcal{Y}}$) in that region. This is greater than ϵ for well chosen δ ($\delta = O(\sqrt{\frac{\epsilon}{dx'}})$), which allows us to conclude and get the bound we want, as $x'/8\sqrt{r\delta} = O(1/\epsilon^{1/4})$. □

6. Upper Bound

We show an upper bound on the complexity of any auction that obtains all but at most ϵ of the optimal revenue. This type of result is useful because it shows that we can actually construct an auction which comes close to obtaining the maximum revenue. In general, when we are forced to take problems that are naturally more continuous and make them into discrete problems, we rely on one type of trick: *rounding*. This type of approach has seen numerous applications in fields of mechanism design, algorithmic game theory, and learning theory. For example, rounding is frequently used in learning theory to discretize the hypothesis class. However, rounding often comes with certain difficulties. "Sharp edges" can occur, for example, when an algorithm that takes the maximum of many elements suddenly takes a dramatically different element when each of the elements are perturbed only slightly. An example can be seen below:

$$\{\mathbf{0.4998}, 0.4999, 0.5000, 0.5001, 0.5002\} \to \{0.5002, 0.5001, 0.5000, 0.4999, \mathbf{0.4998}\}.$$

Despite each individual element not changing by much, the maximum value moved from the first element to the last element. This could create large difficulties in algorithms with rounding that are sensitive to maxima, so we must be careful to ensure this does not affect our algorithm. In our case, this could be dangerous because slightly perturbing the prices along a boundary between valuations may shift the optimal bundle $x(z)$ to a much cheaper bundle, which could drastically hurt revenue.

In the auction problem, we take the optimal menu (which may have infinite size), and then round each of the items on the menu to rational numbers with fixed denominators, which are then placed on an approximate menu with finite size (because there are a finite number of rational numbers with low denominators). In theory, this type of mechanism should come close to the correct one, because it is "basically" the same as the other mechanism, with values rounded. In order to state our theorem, we first introduce the McAfee-McMillan conditions, which enforce certain "niceness" conditions on a distribution.

Definition 6.1. *A distribution $F \in \Delta([0,1]^n)$ is said to satisfy the* McAfee-McMillan conditions *if it has a differentiable density function f satisfying*

$$(n+1)f(x) + x \cdot \nabla f(x) \geq 0,$$

for every $x = (x_1, \ldots, x_n) \in [0,1]^n$

One way to think about the McAfee-McMillan conditions is that there is no region of space where the densities change very suddenly, in essence, that the gradient of the density function is not large in any particular location. We now use a recently proved theorem:

Theorem 6.1 ([10], Proposition 1)**.** *For every distribution F supported on $[0,1]^2$ that satisfies the McAfee-McMillan hazard condition (for $n = 2$), there exists a revenue-maximizing mechanism $M = (x, p)$ that has no outcome for which both allocation probabilities are in the open interval $(0,1)$. In other words, the bundles $x(z)$ for all values of z are on the boundary of the unit square $[0,1]^2$.*

We are now ready to state our main theorem, and we have the tools necessary to prove it.

Theorem 6.2 ([5], Theorem 4)**.** *There exists $C(\epsilon) = O(1/\epsilon^2)$ such that for every $\epsilon > 0$ and for every distribution $F \in \Delta([0,1]^2)$ satisfying the McAfee-McMillan hazard condition, there exists a mechanism M which with menu-size at most $C(\epsilon)$ such that $\mathrm{Rev}(F, M) > \mathrm{OPT}(M) - \epsilon$.*

Proof of Theorem 6.2. Let M be a revenue-maximizing mechanism for F. Let $\lfloor t \rfloor_\delta$ be the rounding down of t to the nearest integer multiple of δ.

Now, before we construct the rounded mechanism, we make a few observations about the nature of the rounding process. Under a mechanism F' which gives the exact same probabilities of allocation, just with adjusted prices, what conditions to we want to be satisfied? The following is an example of something that could go wrong if we naively round the prices.

Example 6.1. Assume that the menu has two components.

- One item to the buyer with probability 0.5 with price 1.01
- One item to the buyer with probability 0.999 with price 2.

If the buyer's value for the object is 2, the first option gives him -1¢ in expected utility, and the second option gives him -0.2¢ in expected utility, so he will choose the second option.

Instead, imagine we round the menu prices, so that the item with probability 0.5 has price 1 (changed) and the item with probability 0.999 has price 2 (stays the same). Now, note that the buyer's decision changes to buying item 1. The utility function for option 1 is now zero, which is higher than -0.2, and the buyer chooses option 1. Although the change in the buyer's utility is slight, note that the auctioneer's revenue goes from 2 to 1, which is a substantial cut.

The problem in this situation is that the rounding was done without any care. The rounding should be forced to only increase the revenue of the auction (or at least only reduce it slightly), not decrease it arbitrarily. To do this, we will also discount all the prices by a fixed proportion $1-\epsilon$, which we claim will *ceteris paribus* make the more expensive bundles look relatively more attractive by properties of rounding.

Formally, take each item of the form (x_1, x_2, p) from the menu. Now, replace it with $(x_1, x_2, (1-\epsilon) \times \lfloor p \rfloor_{\epsilon^2})$. The key idea is now akin to a respect-for-improvements result; any item which was both much more expensive and more desirable than another item retains these properties after we apply these operations. To show this, it suffices to prove the following.

Lemma 6.3. *Given the function $f(p) = (1-\epsilon) \times \lfloor p \rfloor_{\epsilon^2}$, for any $p, q \geq 0$ such that $p > q + \epsilon$,*

$$f(p) - f(q) < p - q.$$

Proof. Note that $\lfloor p \rfloor_{\epsilon^2} - \lfloor q \rfloor_{\epsilon^2} < p - q + \epsilon^2$. Then, it suffices to show that

$$(1-\epsilon)(p - q + \epsilon^2) < p - q,$$

which follows because $\epsilon^2(1-\epsilon) < \epsilon^2 < \epsilon(p-q)$. □

Now, suppose that p^* is the price of the bundle chosen before applying f to the mechanism, and $f(q^*)$ is the price bundle chosen after applying f to the mechanism. We know that

$$v_p - p^* \geq v_q - q^*,$$
$$v_p - f(p^*) \leq v_q - f(q^*),$$

by utility considerations, where v_p is the value of the bundle with price p^* and v_q is the value of the bundle with price q^*. Subtracting these, we obtain that

$$p^* - q^* \leq f(p^*) - f(q^*),$$

which by the contrapositive of Lemma 6.3 implies that $p^* - q^* < \epsilon$. Now, we can then calculate that $p^* - f(q^*) < 2\epsilon$, implying for any valuation z received by the buyer, the amount he pays only goes down by at most 2ϵ. This is also true in aggregate, so the revenue is been reduced by at most 2ϵ overall.

Finally, we ask the question; how many items are now on the menu? It may appear at first that this requires a very large number of objects, but instead note that it needs at most $4\left(1 + \frac{1}{\epsilon^2}\right)$. Observe that by Theorem 6.1, all of the items either have guaranteed allocations of the first object or the second object, or have guaranteed not allocations of the first object or second object. Within each of these, there is no reason to keep any "dominated" bundle on the menu, so we can simply take the highest probability item with the same price, of which there are at most $\left(1 + \frac{1}{\epsilon^2}\right)$ of after rounding, and we are done. □

With substantially more effort, one can show the following result.

Lemma 6.4 ([6])**.** *The menu-size required for revenue maximization up to additive ϵ when selling n goods is $(\log n/\epsilon)^{O(n)}$*

This implies the following theorem, which neatly summarizes all of our results.

Theorem 6.5 ([5])**.** *Fix $n \geq 2$. There exists $D_n(\epsilon) = \Theta(\log(1/\epsilon))$ such that for every $\epsilon > 0$ it is the case that $D_n(\epsilon)$ is the minimum communication complexity that satisfies the following: for every distribution $F \in \Delta([0,1]^n)$ there exists a mechanism selling n goods such that the deterministic communication complexity of running M is $D_n(\epsilon)$ and such that $\mathrm{REV}(F, M) > \mathrm{OPT}(F) - \epsilon$. This continues to hold even in the independent setting.*

7. Future Directions

The work in this section is original.

Inspired by the trick in the upper bound, where in addition to rounding the prices we had to discount them, we can formulate a relaxed form of the utility-maximizing agent.

Definition 7.1 (ϵ-naive agents)**.** *An ϵ-naive agent whose valuations are z will submit valuations z if for all z' we have*

$$x(z) \cdot z - p(z) \geq x(z') \cdot z - p(z') - \epsilon.$$

(Contrast this to the standard definition of incentive compatibility where $\epsilon = 0$.)

This ϵ might be thought of as "a person won't cheat the system to save a tenth of a cent". In practice, such a constraint may be achieved by listing prices rounded to the nearest cent or dollar. Now, note that if we round prices *up* to the nearest multiple of ϵ, the revenues will only increase, and an ϵ-naive agent will still choose the same bundles. By simply rounding all prices by ϵ_2, this gives us the following result.

Theorem 7.1. *Assume that for an ϵ_1-naive agent whose value distribution is F, there is a mechanism M which attains a revenue r. Then, there exists a mechanism where the prices are all multiples of ϵ_2 which attains a revenue at least r for an $\epsilon_1 + \epsilon_2$-naive agent.*

Note that while this theorem appears superficially similar to Theorem 6.2, in actuality it is substantially stronger, as it does not assume the McAfee-McMillan conditions that the work in Section 6 assumes.

References

[1] Moshe Babaioff, Yannai A Gonczarowski, and Noam Nisan. The menu-size complexity of revenue approximation. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 869–877. ACM, 2017.

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

[3] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. Mechanism design via optimal transport. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 269–286. ACM, 2013.

[4] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. Strong duality for a multiple-good monopolist. *Econometrica*, 85(3):735–767, 2017.

[5] Yannai A Gonczarowski. Bounding the menu-size of approximately optimal auctions via optimal-transport duality. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 123–131. ACM, 2018.

[6] Sergiu Hart and Noam Nisan. Approximate revenue maximization with multiple items. *Journal of Economic Theory*, 172:313–347, 2017.

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

[8] Gaspard Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'Académie Royale des Sciences de Paris*, 1781.

[9] Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

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