Multi-Item Auctions Defying Intuition?

CONSTANTINOS DASKALAKIS[1]
Massachusetts Institute of Technology

The best way to sell n items to a buyer who values each of them independently and uniformly randomly in $[c, c+1]$ is to bundle them together, as long as c is large enough. Still, for any c, the grand bundling mechanism is never optimal for large enough n, despite the sharp concentration of the buyer's total value for the items as n grows. Optimal multi-item mechanisms are rife with unintuitive properties, making multi-item generalizations of Myerson's celebrated mechanism a daunting task. We survey recent work on the structure and computational complexity of revenue-optimal multi-item mechanisms, providing structural as well as algorithmic generalizations of Myerson's result to multi-item settings.

Categories and Subject Descriptors: F.0 [**Theory of Computation**]: General

General Terms: Algorithms, Economics, Theory

Additional Key Words and Phrases: Auctions, Multidimensional Mechanism Design

1. INTRODUCTION

Optimal mechanism design is a problem with important applications and deep mathematical structure. In its basic formulation, studied in this survey, a seller has n items to sell to m interested buyers. Each buyer knows his own values for the items, but the seller and the other buyers only know a distribution from which these values are assumed to be drawn. The goal is to design a sales procedure, called a *mechanism*, that optimizes the expected revenue of the seller.

The basic version of the problem and its myriad extensions have familiar applications. Here are a few quick ones: When auction-houses sell items, this is the problem that they face. This is also the problem that governments face when auctioning a valuable public resource such as wireless spectrum. Finally, the problem arises every millisecond as auctions are used in sponsored search and the allocation of banner advertisements.

When it comes to selling a single item, optimal mechanism design is really well understood. Building on Myerson's celebrated work [Myerson 1981], it has been studied intensely for decades in both Economics and Computer Science. This research has revealed surprisingly elegant structure in the optimal mechanism, as well as robustness to the details of the distributions, and has had a deep impact in the broader field of mechanism design.

While all this progress has been taking place on the single-item front, the multi-item version of the problem has remained poorly understood. Despite substantial research effort, it is not even known how to optimally sell two items to one buyer. On the contrary, multi-item auctions appear to have very rich structure, often

[1]Supported by a Microsoft Research faculty fellowship, NSF Awards CCF-0953960 (CAREER) and CCF-110149, and ONR grant N00014-12-1-0999.

Author's address: `costis@csail.mit.edu`

exhibiting unintuitive properties.

In this survey, we present recent progress on the structure and computation of optimal multi-item auctions. Our survey will focus on the philosophy and intuition behind the results, but we will also give substantial technical detail. Our intention is not to present a complete account of results on multi-item mechanisms, but to motivate what we view as a fresh perspective on the problem.

Structure. In Section 2, we take a brief tour of the wondrous land of single-item problems, where we discuss the surprising simplicity of optimal mechanisms. Then, in Section 3, we present multi-item examples showing various ways in which the simplicity of optimal single-item mechanisms fails to generalize. A few of these examples are particularly striking as they illustrate quite unintuitive properties of optimal multi-item mechanisms. The structure of these mechanisms appears so rich that we go one step back, in Section 4, discussing approaches that may be able to accommodate this richness. Then, in Section 5, we present a duality based approach to the multi-item problem, showing how it can be used to characterize single-bidder mechanisms. We also show how this framework can be used to demystify the examples of Section 3, which all pertain to a single bidder. In Section 6, we turn to computation, presenting computationally efficient algorithms for the multi-item multi-bidder problem. As a byproduct, these algorithmic results offer a crisp characterization of the structure of optimal multi-item multi-bidder mechanisms. Finally, in Section 7, we wrap up with a short summary and future directions.

2. THE WONDROUS MYERSON-LAND

Consider the task of selling an item to a single buyer with the goal of maximizing the seller's revenue. The following is a well-known fact.

Fact 1 [Myerson 1981; Riley and Zeckhauser 1983]. *The optimal way to sell one item to a buyer whose value for the item is drawn from some known distribution F is a take-it-or-leave-it offer of the item at some price in* $\arg\max\{x \cdot (1 - F(x))\}$.

Example 1. *The optimal way to sell one item to a buyer whose value for the item is uniformly distributed in* $[0, 1]$ *is to price the item at* 0.5. *The expected revenue is* 0.25.

While it is perhaps intuitive that this should be true, it still surprising that, among all possible communication protocols that the seller and buyer could engage in, the optimal one would require minimal communication and involve no randomization at all.

In this light, it is even more surprising that the optimal auction would maintain its simplicity when multiple buyers are involved.

Fact 2 [Myerson 1981]. *When one item is sold to m buyers whose values for the item are i.i.d. from a known, regular distribution F,*[2] *the optimal mechanism is a second price auction with reservation price* $\arg\max\{x \cdot (1 - F(x))\}$.

[2]A distribution F with density f is called *regular* iff $x - \frac{1-F(x)}{f(x)}$ is monotone increasing.

There are several reasons why this fact is surprising, besides the simplicity of the optimal mechanism:

(1) The mechanism is *deterministic*, namely the auctioneer does not need access to a random generator to implement the allocation and pricing.
(2) The mechanism requires only *one round of communication*, namely the bidders submit bids to the auctioneer who then decides the outcome without any further exchanges with them, except to announce the outcome.
(3) The mechanism is *dominant strategy truthful (DST)*, but it is optimal among the larger class of Bayesian Incentive Compatible (BIC) mechanisms.[3]

In fact, these properties carry over to settings where bidder values are not necessarily i.i.d. and their distributions are not necessarily regular.

FACT 3 [MYERSON 1981]. *When one item is sold to m buyers whose values for the item are independently drawn from known distributions, the optimal mechanism is a* virtual welfare maximizer*, namely:*

(1) Bidders are asked to report bids for the item: $b_1, \ldots, b_m$.
(2) Bids are transformed into what are called "ironed virtual bids," $h_1(b_1), \ldots, h_m(b_m)$, where each $h_i(\cdot)$ depends on the corresponding bidder's value distribution (but not on the other bidders' distributions and not even m).[4]
(3) The item is allocated to the bidder with the highest ironed virtual bid, with some lexicographic tie-breaking.
(4) The winner of the item is charged his threshold bid, namely the smallest bid he could place and still win the item.

Facts 1-3 are both surprising and powerful, providing simple, yet sharp and versatile machinery for revenue optimization in single-item and, more broadly, single-dimensional environments.[5] Importantly, they have provided solid foundation for a tremendous literature that has brought tools from approximation algorithms and probability theory into mechanism design. Building on the shoulders of Myerson, this literature strives to understand how to make the theory robust, further improving the simplicity of mechanisms and reducing their dependence on the details of the bidders' distributions, both at a quantifiable loss in revenue. See, e.g., [Hartline 2013; Chawla and Sivan 2014; Roughgarden 2015] for recent surveys of this work.

[3] A mechanism is called *Dominant Strategy Truthful* iff it is in the best interest of every bidder to truthfully report their value to the mechanism, regardless of what the other bidders report. Bayesian Incentive Compatible mechanisms are a broader class, but we postpone their definition, as it is slightly technical and not important for our discussion right now. See Definition 6.

[4] The precise functional form of the h_i's is not important for this survey.

[5] In a single-dimensional environment, the seller can provide service to several buyers, subject to constraints on which buyers can receive service simultaneously. Each buyer has a value for receiving service, which is distributed according to some distribution known to the seller and the other buyers. Single-dimensional environments clearly generalize single-item environments, where only one buyer can be served the item.

3. AUCTIONS DEFYING INTUITION

Despite remarkable progress on the single-item front over the past few decades, revenue optimization in multi-item settings has remained poorly understood. We do not even have a sharp characterization of optimal two-item mechanisms, even when there is a single buyer. On the contrary, multi-item mechanisms exhibit such rich structure that it is difficult to imagine what a generalization of Myerson's results could look like. Or, better said, the generalizations that we can imagine can be shot down via simple examples.

To illustrate the richness of multi-item mechanisms, let us consider some simple multi-item settings and their corresponding optimal mechanisms. All our examples will involve a seller with n items and a single *additive buyer*. Such a buyer is characterized by a private vector $(v_1, \ldots, v_n)$ of values for the items and derives utility $\sum_{i\in S} v_i - p$, whenever he pays p to get the items in set $S \subseteq [n]$. If S and p are random his utility is $\mathbb{E}[\sum_{i\in S} v_i - p]$.

3.1 Bundling

When it comes to multiple items, a natural question to ask is whether we can use Myerson's technology to design optimal mechanisms, and indeed what exactly it is that we should be selling. The following simple example illustrates that we may need to bundle items, even when there is a priori no interaction between them.

Example 2. *Suppose $n = 2$ and the buyer's values are i.i.d., uniform in $\{1, 2\}$. Since the buyer is additive, and his values for the items are independent, getting one of the two items will not affect his marginal value for getting the other item as well. Since there is no interaction between the item values, it is natural to expect that the optimal mechanism should sell the two items separately. By Fact 1, this would mean pricing each item at 1 and letting the buyer decide which of them to buy. Simple calculations show that the expected revenue of this mechanism is 2.*

Interestingly, there is a flaw in this logic. While it is true that item-values do not "interact with each other," we may still want to capitalize on the fact that the buyer's average item-value is better concentrated than his value for a specific item. Indeed, it is better to only offer the bundle $\{1, 2\}$ of both the items at price 3. Given that $\Pr[\sum_i v_i \geq 3] = 3/4$, the expected revenue of the seller is now $9/4 > 2$. It can be shown that this is the optimal mechanism [Daskalakis et al. 2014].

Example 2 illustrates the following.

Fact 4. *Optimal multi-item mechanisms may require bundling, even when there is a single additive buyer with independent values for the items.*

Our intuition for the effectiveness of bundling in Example 2 appealed to the concentration of the buyer's surplus (i.e. total value for both items). There was still a flaw in our logic, however, and this time a less tangible one: Why is the surplus the right benchmark to compare against? The following result, discussed in more detail in Section 5.8, illustrates a setting where the structure of the optimal mechanism is different in two asymptotic regimes.

Theorem 1 [Daskalakis et al. 2015]. *Consider selling n items to a buyer whose values for the items are i.i.d. uniform in $[c, c+1]$. The following are true:*

(*1*) *For all n, there exists c_0 such that, for all $c > c_0$, the optimal mechanism only offers the bundle of all the items together at some price.*

(*2*) *For all c, there exists n_0 such that, for all $n > n_0$, the optimal mechanism does not only offer the bundle of all items together at some price.*

Part 2 of the above theorem is especially counter-intuitive. As $n \to \infty$, the buyer's average value for the items, $\frac{\sum_i v_i}{n}$, becomes more and more concentrated around its mean, $c + 0.5$. It is clear that the seller cannot hope to extract higher revenue than the buyer's total expected value, $n(c + 0.5)$, and offering the grand bundle for $n(c + 0.5 - \epsilon)$, for the tiniest discount $\epsilon > 0$, would make the buyer accept to purchase it with probability arbitrarily close to 1 as $n \to \infty$. Still, for no n does this intuition materialize, and it never becomes optimal to only sell the grand bundle...

3.2 Randomization

Recall that optimal single-item mechanisms do not require randomization. Our next example illustrates that this is not the case in multi-item settings.

Example 3. *Suppose $n = 2$, and v_1 is distributed uniformly in $\{1, 2\}$ while v_2 is independently distributed uniformly in $\{1, 3\}$. In this example, an optimal deterministic mechanism prices item 1 at 1 and item 2 at 3. Its expected revenue is 2.5.*

Still, we can do better. We can offer the buyer two options: The first is to pay 4 and get both items. The second is to pay 2.5 to get a "lottery ticket" that allocates item 1 with probability 1 and item 2 with probability 1/2. The expected revenue of this mechanism is 2.625, which can be shown to be optimal [Daskalakis et al. 2014].

So randomization is necessary for revenue maximization. It turns out its effect on the revenue may actually be quite dramatic.

Fact 5. *Optimal multi-item mechanisms may require randomization. The gap between the revenue of the optimal randomized and the optimal deterministic mechanism can be arbitrary large, even when there are two items and a single buyer [Briest et al. 2010; Hart and Nisan 2013].*

3.3 Menu Size Complexity

In our previous examples, we described the optimal mechanism as a menu of options for the buyer to choose from. If the optimal mechanism were guaranteed to be deterministic, describing it as a menu would require a bounded number of options, as there is a finite number of possible bundles that the mechanism may offer. Given Fact 5, however, it becomes unclear how to specify the optimal mechanism. The following example illustrates that representing it as an explicit menu may be infeasible.

Example 4 [Daskalakis et al. 2013]. *Suppose $n = 2$, and v_1 is distributed according to the Beta distribution with parameters $(3, 3)$ while v_2 is independently distributed according to the Beta distribution with parameters $(3, 4)$.[6] Then the*

[6]The *Beta distribution* with parameters (α, β) is distributed in $[0, 1]$ according to the density function $f(x) \propto x^{\alpha-1}(1 - x)^{\beta-1}$.

optimal mechanism needs to offer uncountably many lotteries.

The alarming feature of Example 4 is that it becomes unclear whether the allocation and price rule of the optimal mechanism can be effectively described via a small number of parameters, even if the buyer's distribution can be.[7] In Section 5.6.2, we will show that the optimal mechanism for Example 4 can actually be described via a small number of parameters. However, this may not be true in general.

3.4 Non-Monotonicity

It seems intuitive that a seller with more valuable items should expect a higher revenue from selling them. One way to quantify this intuition is the following.

Consider two distributions F and G such that F first-order stochastically dominates G, denoted $F \succeq_1 G$. This means that, for all $x \in \mathbb{R}$, $F(x) \leq G(x)$, i.e. F and G can be coupled so that F always samples a value larger than the value sampled by G. It easily follows from Fact 1 that a seller selling an item to some buyer whose value for the item is distributed according to F makes higher revenue than if the buyer's value were distributed according to G.

Surprisingly this fails to hold in multi-item settings!

FACT 6 [HART AND RENY 2012]. *There exist distributions F and G such that $F \succeq_1 G$ but the optimal revenue from selling two items to a buyer whose values are i.i.d. from F is smaller than if they were i.i.d. from G.*

4. CROSSROADS

It is clear from our examples in the previous section that, in the close neighborhood of Myerson's setting, optimal mechanisms exhibit rich structure and may defy our intuition. It is not even clear if they have a finite effective representation. In view of these complications, there are several directions we may want to pursue:

(1) Forget about trying to understand optimal mechanisms and pursue approximations directly. Even though optimal mechanisms are complex, there may still be simple mechanisms that can be shown to guarantee some good fraction of the optimal revenue.

(2) Forget about characterizing the structure of optimal mechanisms and study instead whether they can be computed efficiently.

(3) Develop new machinery to characterize the structure of optimal mechanisms. Despite their apparent complexity and fragility, there may still be a different lens through which they exhibit more structure.

It is a priori dubious whether the approximation approach can lead anywhere. Without understanding the optimal mechanism, how can we possibly establish the approximate optimality of some other mechanism? It is quite surprising then that this approach has actually been quite fruitful:

[7] Of course, in a somewhat perverse way, the buyer's distribution itself "indexes" the optimal mechanism for that distribution. But this does not count as it is unclear if there is an efficient procedure that can implement the mechanism given this description. We will touch upon this point a bit later, when we discuss the computation of optimal mechanisms in Section 6.

—In multi-item settings with additive buyers, [Hart and Nisan 2012; Li and Yao 2013; Babaioff et al. 2014; Yao 2015; Rubinstein and Weinberg 2015] design approximate mechanisms by carefully decomposing the support of the buyers' distributions into regions and, due to lack of a better benchmark, using expected welfare as an obvious upper bound to the optimal revenue, competing against this stronger benchmark in some of the regions. Here is a very interesting and clean result that they obtain for the setting of the previous section.

THEOREM 2 [BABAIOFF ET AL. 2014]. *When n items are sold to a buyer whose values for the items are independent, a mechanism that either only prices individual items or only prices the bundle of all the items obtains at least* $1/6$*-th of the optimal revenue.*

—In multi-item settings with unit-demand buyers,[8] [Chawla et al. 2007; Chawla et al. 2010] upper bound the optimal revenue by the optimal revenue in a related single-dimensional setting, and define sequential posted price mechanisms for the multi-item setting competing against this stronger upper bound.

These approximation results provide simple ways through which a constant fraction of the optimal revenue can be attained, bypassing the difficulties coming from our lack of understanding of optimal multi-item mechanisms. Moreover, they help us understand the tradeoffs between simplicity, optimality and generality of mechanisms. For example, Theorem 2 tells us that, if we are willing to sacrifice generality in the model and 5/6-ths of the revenue, a very simple mechanism will work for us. Still such approximation results only apply to restricted settings, e.g. they typically assume independence of the distributions across items, and it is unclear how to extend them to broader settings: correlation among items, more complex buyer valuations, and more complex allocation constraints.

Ultimately, it is our belief that a cohesive theory of optimal multi-item mechanisms cannot be obtained through disparate approximation results applying to different multi-item environments. And even where these approximation results do apply they may still not provide fine-tuned insight into the salient features of the setting responsible for revenue. As a simple example, the non-monotonicity of revenue (Fact 6) is not foreseeable from Theorem 2 alone. On the contrary, the theorem hints towards the incorrect conclusion.

In this light, over the past several years we have pursued the challenge of characterizing optimal multi-item mechanisms from both an algorithmic and a structural perspective. In the next two sections, we give a flavor of our progress on these fronts. In both sections, we discuss our philosophy as well as give an overview of results and techniques. We note that our goal is not the coverage of all results in the literature, but mostly the philosophy behind them. So we will focus on our philosophy and a biased sample of primarily our own results.

[8]These are buyers who only want to purchase one item.

5. THE STRUCTURE OF MULTI-ITEM MECHANISMS

5.1 Philosophy

In Sections 3.1–3.4 we gave several multi-item examples, making claims about their optimal mechanisms. Taking a step back, how did we go about proving these claims? More broadly, how is the optimality of some object, such as a mechanism, established? Two principled approaches for doing this are the following. The first entails formulating the problem at hand as a convex minimization/concave maximization problem, and then showing that the object whose optimality we want to establish satisfies first-order conditions. The second is to develop a dual formulation and identify a solution to the dual that matches the value attained by the object whose optimality is to be established in the primal.

In single-dimensional environments, Facts 1-3 can be shown using the first approach. Through a chain of deductions, [Myerson 1981] expresses the expected revenue of a mechanism as an expected (virtual) welfare quantity, which can be optimized in a point-wise manner.

In multi-item environments, it is not clear how to generalize Myerson's approach, except under significant restrictions on the value distributions; see [Rochet and Stole 2003; Manelli and Vincent 2007]. This motivates the second approach, which we have pursued in the context of a single, additive buyer in [Daskalakis et al. 2013; 2015]. We give a flavor of our approach with examples in the following sections.

5.2 Setting

In this section, we restrict our attention to a seller with n items and an additive buyer whose values for these items are jointly distributed according to some distribution F. For simplicity, we assume that F is supported on some box $X = \prod_i [\underline{x}_i, \overline{x}_i] \subset \mathbb{R}^n_+$, and is continuously differentiable with bounded partial derivatives. We call X the *typeset* of the buyer, and its elements the buyer's possible *types*. If a buyer of type x is allocated a subset S of the items and pays price t for it, his utility is $\sum_{i \in S} x_i - t$. Our buyers are also risk-neutral, so if S and t are random, then their utility is $\mathbb{E}_{S,t}[\sum_{i \in S} x_i - t]$.

The seller only knows F, but the buyer knows his realized type, and the seller's goal is to design a mechanism to optimize her expected revenue. By the revelation principle the optimal mechanism can be described in the form of a menu whose entries are lotteries. Each lottery specifies a vector of probabilities $p_1, \ldots, p_n$ and a price t. If purchased, it will allocate each item i independently with probability p_i. Given a menu of lotteries, the buyer will choose the lottery that optimizes his utility, given his type. If all lotteries give him negative utility, then he will not purchase any of them.

5.3 Convex Optimization Formulation

To develop our duality framework we start with expressing our mechanism design problem as a convex optimization problem. Here we have a few options.

(1) Represent the mechanism as a menu of lotteries for the buyer to choose from. There are a few reasons why we dislike this approach. First, the menu of the optimal mechanism may be uncountable in size, as Example 4 illustrates. So we would have to represent it as a continuous set. Second, given some representation of the menu, it is cumbersome to express the expected revenue resulting from this menu, as such an expression would have to incorporate the buyer's optimization over lotteries in the menu. Finally, each lottery in the menu is multi-parametric, comprising n allocation probabilities as well as a price.

(2) Represent the mechanism as a menu, but also keep track of which lottery in the menu each possible type of buyer will purchase. In this case, a mechanism can be represented as a pair of functions: `(i)` the *allocation function* $\mathcal{P} : X \to [0,1]^n$ specifying the allocation probabilities of the lottery that each type will purchase, if any; and `(ii)` the *price function* $\mathcal{T} : X \to \mathbb{R}$ specifying the price that each type will pay for the purchased lottery, if any.

Now, the representation of the mechanism is simpler, and we can easily express its expected revenue as follows:

$$\int_X \mathcal{T}(x)dF(x). \tag{1}$$

Of course, we need to add some consistency constraints to make sure that our modeling is faithful:

$$\forall x, x' \in X : x \cdot \mathcal{P}(x) - \mathcal{T}(x) \geq x \cdot \mathcal{P}(x') - \mathcal{T}(x'); \tag{2}$$

$$\forall x \in X : x \cdot \mathcal{P}(x) - \mathcal{T}(x) \geq 0. \tag{3}$$

Constraint (2) expresses that no type prefers a different lottery to the one we maintain for this type, while Constraint (3) expresses that each type will actually buy this lottery.

Finding the pair of functions $(\mathcal{P}, \mathcal{T})$ optimizing (1) subject to the constraints (2) and (3) was the approach taken by [Myerson 1981] and is quite standard. In the multi-item setting, however, this representation is still cumbersome to work with as, besides having an n-variate input, the compound function $(\mathcal{P}, \mathcal{T})$ also has an $(n+1)$-dimensional output.

(3) Given the complexity of the standard representation, we decide to optimize over mechanisms indirectly. Rather than optimizing revenue in terms of the mechanism's allocation and price functions, we want to explore whether we can optimize revenue in terms of the buyer's *utility* from participating in the mechanism.

Indeed, facing some mechanism, a buyer of type x will decide to buy some lottery (p_x, t_x), thereby enjoying utility $p_x \cdot x - t_x$ from his decision.[9] So, every mechanism induces a function $u : X \to \mathbb{R}$, where $u(x)$ expresses the utility of the buyer when his realized type is x and he buys his favorite lottery in the mechanism, if any. Our goal is to optimize over mechanisms indirectly by optimizing over u's, which raises two questions:

(1) Given a function $u : X \to \mathbb{R}$, can we recognize whether there is a mechanism inducing this function u as the utility of the buyer?

(2) If u *is* induced by a mechanism, is there enough information in u to uniquely specify the expected revenue of whatever mechanism induces u, and can we get our hands on a mechanism that induces u?

[9]Let us assume that the lottery $(\vec{0}, 0)$ is always in our menu to account for the possibility that the buyer may derive negative utility from all lotteries in the menu and hence decide to buy none.

A priori it is unclear whether u is informative enough about the mechanism(s) that induce(s) it. In fact, it appears that it is losing information about these mechanisms. Nevertheless, the answer to each of the two questions above is actually "yes," due to the following theorem by Rochet.

THEOREM 3 [ROCHET 1987]. *Function $u : X \to \mathbb{R}$ is induced by a mechanism iff u is 1-Lipschitz continuous with respect to the ℓ_1-norm, non-decreasing, convex and non-negative. Moreover, if these conditions are met then $\nabla u(x)$ exists almost everywhere in X, and wherever it exists:*

—*$\nabla u(x)$ are the allocation probabilities of the lottery purchased by type x, and*
—*$\nabla u(x) \cdot x - u(x)$ is the price of the lottery purchased by type x*

in any mechanism inducing u.

The theorem follows by combining Constraints (2) and (3). For a concise derivation of the theorem, please refer to Lecture 21 of [Daskalakis 2015]. In order to ground it to our experience, let us plot the utility induced by the optimal mechanism in Example 1. The utility is shown in Figure 1. We can verify that it satisfies the conditions of the theorem, and its derivatives contain information about the allocation function of the mechanism.

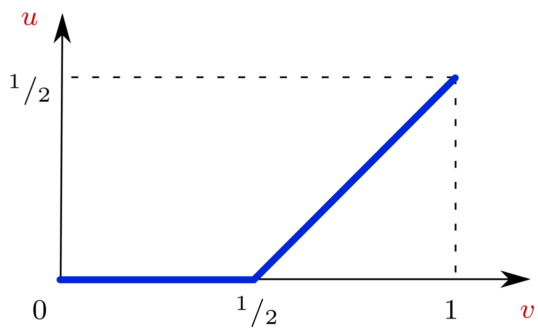

Fig. 1. The utility induced by a take-it-or-leave it offer of an item at 0.5.

With Theorem 3, we can formulate our mechanism design problem as follows:

$$
\begin{aligned}
& \sup \int_X (\nabla u(x) \cdot x - u(x)) dF(x) \\
s.t.\quad & |u(x) - u(y)| \leq |x - y|_1, \forall x, y \in X \\
& u : \text{ non-decreasing} \\
& u : \text{ convex} \\
& u(x) \geq 0, \forall x \in X.
\end{aligned}
$$

Note that our formulation aims at optimizing the expected price paid by the buyer, which is given by $\nabla u(x) \cdot x - u(x)$ when his type is x, subject to the constraints on the function u. Since $\nabla u(x)$ may only be undefined at a measure zero subset of X (given that F has no atoms), we can take "$\nabla u(x)$" to mean anything when it is undefined, and this will not affect our revenue. For concreteness, from now on, whenever we write $\nabla u(x)$, we will mean any subgradient of u at x (which will exist as u is convex).

The advantage of our new formulation is that the variable u is a scalar function of the buyer's type. The downside is that the objective function is cumbersome and the constraints, especially convexity, are difficult to handle. We can eliminate the first issue with a little massaging of the objective, using the divergence theorem. Our objective can equivalently be written as follows, where f denotes the density of F and $\hat{\eta}(x)$ denotes the outer unit normal vector at point $x \in \partial X$ of the boundary of X—see Lecture 21 of [Daskalakis 2015] for a concise derivation.

$$\int_X (\nabla u(x) \cdot x - u(x))dF(x) \equiv \tag{4}$$

$$\int_{\partial X} u(x)f(x)(x \cdot \hat{\eta}(x))dx - \int_X u(x)(\nabla f(x) \cdot x + (n+1)f(x))dx. \tag{5}$$

Our massaged objective is linear in our variable u. Indeed, it can be viewed as the "expectation" of u with respect to the signed measure μ with the following density:[10]

$$f(x)(x \cdot \hat{\eta}(x))\mathbf{1}_{x \in \partial X} - (\nabla f(x) \cdot x + (n+1)f(x)).$$

Equipped with the above definition, we can express our mechanism design problem as the following convex optimization problem.

$$(P): \quad \begin{aligned} & \sup \int_X u(x)d\mu(x) \\ s.t. \quad & |u(x) - u(y)| \le |x - y|_1, \forall x, y \in X && (6) \\ & u: \text{ non-decreasing} && (7) \\ & u: \text{ convex} && (8) \\ & u(x) \ge 0, \forall x \in X. && (9) \end{aligned}$$

Balancing μ. For $u(x) = 1$, integral 4 becomes -1. Hence, so does integral 5, and therefore $\int_X d\mu = -1$. So $\mu(X) = -1$. It is convenient to have μ balanced, namely satisfy $\mu(X) = 0$. So we add an atom of $+1$ at point $\underline{x} = (\underline{x}_1, \ldots, \underline{x}_n)$ to make this happen. Accordingly, from now on, μ is in fact the measure with the following density:

$$\mathbf{1}_{x=\underline{x}} + f(x)(x \cdot \hat{\eta}(x))\mathbf{1}_{x \in \partial X} - (\nabla f(x) \cdot x + (n+1)f(x)). \tag{10}$$

[10] To be formal here, we should not define μ through its density. It is more accurate to define μ using Riesz's representation theorem, as the unique Radon measure such that the bounded linear functional (5) of bounded continuous functions u equals $\int_X u d\mu$. Accordingly, all measures in our duality framework will be Radon measures. Having said that, we invite the reader to forget about Radon measures and Riesz's theorem for the remainder of this survey. Formally, whenever we talk about measures we mean Radon measures, and μ is as defined by Riesz's representation theorem.

5.4 Interpretation of Convex Formulation

Figure 2 shows μ for the single-item setting of Example 1, while Figure 3 shows μ for the setting of Example 4. As measure μ is now two-dimensional we have color-coded it. The dashed curve separates $X_+ \setminus \{\vec{0}\}$ from X_-, where X_+ and X_- are subsets of X where μ is positive and negative respectively. Notice the atom of $+1$ measure at point $\vec{0}$ in both figures.

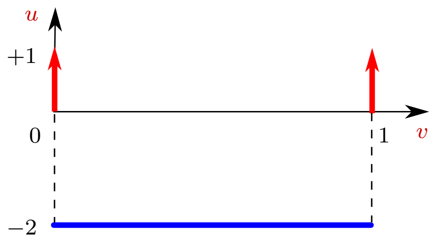

Fig. 2. Measure μ in Example 1.

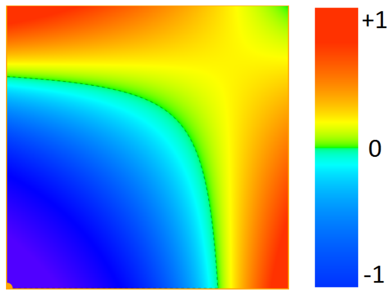

Fig. 3. Measure μ in Example 4. The dashed curve separates $X_+ \setminus \{\vec{0}\}$ from X_-. In particular, $X_+ \setminus \{\vec{0}\}$ is an increasing (ie upper-closed) subset of X.

Formulation (P) is asking us to maximize $\int_X u d\mu$. So we would like to make u as large as possible where μ is positive and make it as small as possible where μ is negative. However, Constraints 6–9 impose interesting tradeoffs on how we should implement this optimally, as u should be continuous, not too steep, and convex-nondecreasing.

For instance, the optimal u against the measure of Figure 2 turns out to be the one shown in Figure 1. How do we know this? It is not completely obvious, but we know it should be true from Fact 1/Example 1.

In the multi-item setting, what we need is new machinery that will allow us to identify optimal such tradeoffs. As we have already discussed, however, we know of no direct way to find these tradeoffs by studying formulation (P) directly. Our approach is instead to develop a dual optimization problem that will hopefully provide insight about the optimal solution to (P).

5.5 Duality: Building Intuition

It is a priori not clear whether (P) has a strong dual problem, i.e. a minimization problem whose optimal value matches that of the optimal value of (P). The reason is that it is an infinite-dimensional program, and we are not aware of infinite-dimensional programming tools that can handle our constraints [Luenberger 1968; Anderson and Nash 1987].

We will show that a strong dual actually does exist in Section 5.7. Prior to that let us build some intuition though, grounding it to our experience in the finite-dimensional world. In this section, we will make an analogy to min-cost perfect matching duality that will lead us to formulate a weak dual of (P), upper-bounding but not necessarily matching the value of (P). But first let us massage (P) to the following equivalent formulation:

$$
(P): \quad \begin{aligned} & \sup \int_X u(x) d\mu(x) & \\ s.t. \quad & u(x) - u(y) \leq |(x-y)_+|_1, \forall x, y \in X & (11) \\ & u: \text{ convex} & (12) \\ & u(\vec{0}) = 0 & (13) \end{aligned}
$$

where we denote by $|(x-y)_+|_1 = \sum_i \max(0, x_i - y_i)$. Notice that Constraint (11) already implies that u must be non-decreasing as well as 1-Lipschitz with respect to the ℓ_1-norm. Combined with (13), it also implies non-negativity. So, if anything, our new constraints might have restricted the set of functions that we are optimizing over. But, it is also easy to see that Constraints (6) and (7) imply Constraint (11). Moreover, Constraint (13) could have been added to the original formulation without changing the optimal value, given that $\mu(X) = 0$. So the two formulations are actually equivalent.

Next, let us write μ as the difference $\mu_+ - \mu_-$ of two non-negative measures μ_+ and μ_-, so that

$$\int_X u d\mu = \int_X u d\mu_+ - \int_X u d\mu_-$$

for all measurable u. Given that $\mu(X) = 0$, it follows that $\mu_+(X) = \mu_-(X)$. Moreover, let X_+ and X_- be a partition of X so that μ_+ is supported on X_+ and μ_- is supported on X_-. With this notation, let us consider the following relaxation of (P):

$$
(P'): \quad \begin{aligned} & \sup \int_X u d\mu_+ - \int_X u d\mu_- \\ s.t. \quad & u(x) - u(y) \leq |(x-y)_+|_1, \forall x \in X_+, y \in X_- \\ & u(\vec{0}) = 0 \end{aligned}
$$

where we have dropped the convexity constraint, and only maintain Constraint 11 for $x \in X_+$ and $y \in X_-$. It is clear that any feasible solution to (P) is also a feasible solution to (P') and, therefore, the optimum of (P') upper bounds the optimum of (P).

We are thus ready to employ our finite-dimensional intuition. Suppose temporarily that sets X_+ and X_- were finite, and measures μ_+, μ_- were uniform over X_+ and X_- respectively. In this case (P') becomes a problem of assigning potential values to the nodes of the complete bipartite graph $(X_+, X_-, X_+ \times X_-)$ with the goal of optimizing the total potential (gaining the potential of nodes in X_+ and losing the potential of nodes in X_-) subject to the constraint that for every $x \in X_+$ and $y \in X_-$ the difference in potential between x and y cannot exceed $|(x-y)_+|_1$. It is well-known that the dual of this problem is a min-cost perfect matching problem on the same graph, where the weights are as in Figure 4.

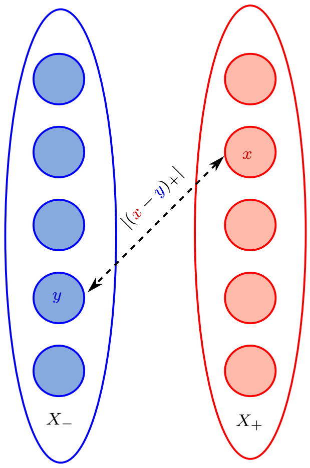

Fig. 4. Min-Cost Perfect Matching.

Going back to the infinite-dimensional problem, a perfect matching should correspond to a coupling of the measures μ_+ and μ_-, defined formally as follows.

DEFINITION 1. *A* coupling *between two non-negative measures μ_1 and μ_2 defined over some $S \subseteq X$ and satisfying $\mu_1(S) = \mu_2(S)$ is a non-negative measure γ over $S \times S$ such that, for all measurable subsets $S' \subseteq S$, it holds that:*

$$\int_{S' \times S} d\gamma = \mu_1(S')\mu_2(S) \quad \text{and} \quad \int_{S \times S'} d\gamma = \mu_1(S)\mu_2(S').$$

We denote by $\Gamma(\mu_1, \mu_2)$ the set of all couplings between μ_1 and μ_2.

Given the above definition and our finite-dimensional intuition, it makes sense to propose the following as a dual to (P'):

$$(D'): \quad \inf \int_{X \times X} |(x-y)_+|_1 d\gamma(x,y)$$
$$s.t. \quad \gamma \in \Gamma(\mu_+, \mu_-)$$

It is indeed quite straightforward to establish that the optimum of (D') upper-bounds the optimum of (P').

LEMMA 1 [DASKALAKIS ET AL. 2013]. *(D') is a weak dual of (P').*

PROOF (LEMMA 1): For any feasible solution u of (P') and γ of (D') we have the following:

$$\int_X u d\mu = \int_X u d(\mu_+ - \mu_-) = \int_{X \times X} (u(x) - u(y)) d\gamma(x,y) \tag{14}$$
$$\leq \int_{X \times X} |(x-y)_+| d\gamma(x,y), \tag{15}$$

where the second equality follows from the feasibility of γ in (D') and the inequality follows from the feasibility of u in (P'). $\square$

5.6 Applications of Weak Duality

So far, we have formulated revenue optimization as a convex optimization problem (P), defined a relaxation (P') of (P) and identified a weak dual (D') of (P') and therefore of (P). Namely, here is where we stand:

$$\mathrm{OPT}(P) \leq \mathrm{OPT}(P') \leq \mathrm{OPT}(D').$$

Given that (D') is only a weak dual of (P), it is unclear how to use it to certify optimality of mechanisms. Indeed, it could be that u is an optimal solution to (P), but there is no solution γ of (D') that matches it in value of the objective. So what is the point of (D')?

The mismatch between the values of (P) and (D') motivates the development of a strong dual of (P) in the next section. Nevertheless, our experience shows that (D') often gives a simple heuristic that can be used to certify optimality of mechanisms. In the remainder of this section, we illustrate how (D') can be applied for this purpose.

Suppose that u is the utility function of some mechanism that we want to show is optimal for some distribution F. So we want to show that u is an optimal solution to (P), where μ is the signed measure derived from F according to (10). Even though (D') is only a weak dual of (P) we could still be optimistic, trying to find a feasible solution γ of (D') that achieves the same objective value as that achieved by u in (P).

What does γ need to satisfy for this to happen? Inspecting the proof of Lemma 1 (which also applies to feasible solutions of (P) and (D')) we get the following sufficient condition:

$$\gamma(x,y)\text{-almost everywhere:} \;\; u(x) - u(y) = |(x-y)_+|. \;\; ^{11}$$

This condition would force the only inequality in the proof to be tight. Given that the gradient of u, wherever defined, corresponds to allocation probabilities, the following condition is also sufficient:

$\gamma(x,y)$-almost everywhere:

$$\begin{pmatrix}\nabla u \text{ exists on all points} \\ \text{of the segment } (x,y)\end{pmatrix} \;\wedge\; \forall i : \begin{cases} x_i > y_i \implies & \text{all types on segment } (x,y) \text{ receive item } i \text{ with probability } 1 \\ x_i < y_i \implies & \text{all types on segment } (x,y) \text{ receive item } i \text{ with probability } 0 \end{cases} \tag{16}$$

Equipped with sufficient condition (16), we exhibit how it can be used to show optimality of mechanisms. We start with a simple example in Section 5.6.1, pro-

[11] What this means is that the set of points $(x,y) \in X \times X$ where the equality fails to hold has measure 0 under γ.

ceeding to a more complex example in Section 5.6.2. It will be instructive to think of our couplings γ as routing infinitesimal amounts of flow between pairs of points in X, and Condition (16) as restricting the directions of these flows.

5.6.1 *Certifying Optimality of Mechanisms Using Duality*

Example 5. *Suppose* 2 *items are sold to a buyer whose values for the items are i.i.d., uniform* $[0, 1]$.

Claim 1 [Manelli and Vincent 2006]. *The optimal mechanism in Example 5 is to price each individual item at 2/3, and the bundle of both items at* $\frac{4-\sqrt{2}}{3}$.

Proof (Claim 1): First, let us compute the measure μ induced by the uniform distribution over $X = [0, 1]^2$ according to (10). The resulting measure comprises a total of -3 surface measure distributed uniformly over $[0, 1]^2$, a total of $+2$ single-dimensional measure distributed uniformly over the top and the right edge of $[0, 1]^2$, and an atom of $+1$ at the origin.

In Figure 5, we partition $[0, 1]^2$ into four regions R_0, R_1, R_2, R_3 corresponding to the subsets of types that will purchase each lottery in the menu, or no lottery at all. (The tie-breaking at the boundaries between regions is unimportant as it corresponds to a measure 0 set of types.) We also depict μ and write down the utility function restricted to each region.

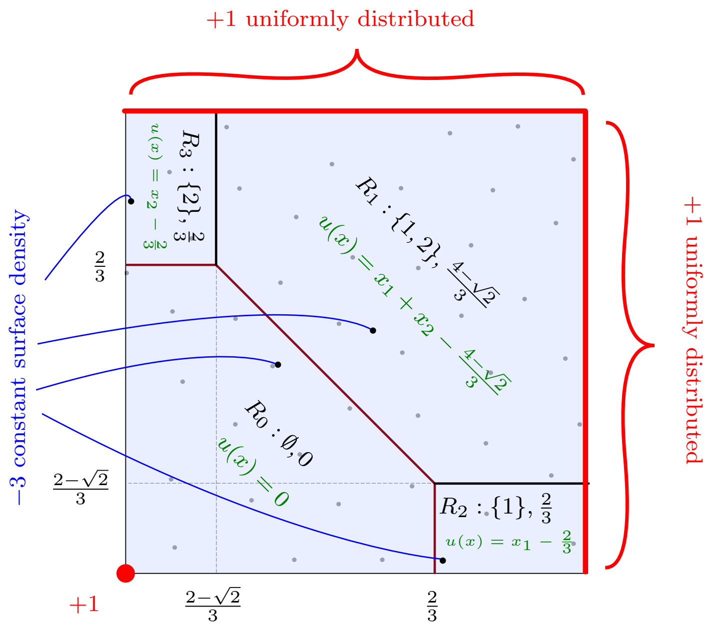

Fig. 5. Measure μ for two i.i.d. uniform $[0, 1]$ items, and the partition of the typeset induced by the mechanism that prices each individual item at 2/3 and the bundle of both items at $\frac{4-\sqrt{2}}{3}$. We also depict the utility of the buyer depending on which region his type falls into.

In particular, region R_0 corresponds to the types that do not want to purchase anything. In this region, there is -1 total surface measure uniformly distributed, as well as an atom of $+1$ total measure at the origin. Region R_1 corresponds to the types that will purchase bundle $\{1, 2\}$ at price $\frac{4-\sqrt{2}}{3}$. There is a total of $-\frac{2+2\sqrt{2}}{3}$ surface measure uniformly distributed here, as well as a total of $+\frac{2+2\sqrt{2}}{3}$ single-dimensional measure spread uniformly on the top and right edges of this region. Finally, regions R_2 and R_3 correspond to types that will purchase items 1 and 2 respectively at price $2/3$. Each of these have a total of $-\frac{2-\sqrt{2}}{3}$ surface measure uniformly distributed, as well as a $+\frac{2-\sqrt{2}}{3}$ single-dimensional measure uniformly distributed on their right and, respectively, top edges.

Our goal is to find a coupling γ between μ_+ and μ_- satisfying Condition (16). Given that our sufficient condition is sensitive about the existence of ∇u, we should consider couplings of μ_+ and μ_- in each region separately. This is conceivable as μ is balanced in each region, namely $\mu_+(R_i) = \mu_-(R_i)$, for all $i = 0, 1, 2, 3$. In fact, our path is cut out for us given the functional form of u in each region and the form of Condition (16). In particular, we are seeking a coupling between μ_+ and μ_- so that

—In region R_0 we are only allowed to transport measure in north-east directions, as both items are allocated with probability 0.

—In region R_1 we are only allowed to transport measure in south-west directions, as both items are allocated with probability 1.

—In region R_2 we are only allowed to transport measure in north-west directions, as item 1 is allocated with probability 1 and item 2 is allocated with probability 0.

—In region R_3 we are only allowed to transport measure in south-east directions, as item 2 is allocated with probability 1 and item 1 is allocated with probability 0.

In fact, we will be more optimistic, restricting our transports to be westward and southward in regions R_2 and R_3 respectively. All in all, we are seeking a coupling of μ_+ and μ_- that pushes measure in the directions shown in Figure 6 in each region.

Now it is easy to verify that the way our regions and measure μ are set up, it is possible to couple μ_+ and μ_- where all transports take place according to the figure. So Condition (16) is satisfied, and the resulting coupling certifies the optimality of u. $\square$

We refer the reader to [Giannakopoulos and Koutsoupias 2014] for an application of the afore-described approach to $n = 3, \ldots, 6$ i.i.d. uniform $[0, 1]$ items. While written in a slightly different language, their proof establishes the existence of solutions to (D') matching the value achieved by the optimal mechanism in (P). It still remains an interesting open problem to determine the optimal mechanism for $n > 6$. We conjecture that using (D') remains sufficient, but it becomes analytically challenging to define the transports in high dimensions.

5.6.2 *Reverse-Engineering Optimal Mechanisms Using Duality.* In Section 5.6.1, we started with a conjectured optimal mechanism. Given the mechanism, we partitioned the typeset into regions, depending on what lottery each type will purchase. We then used Condition 16 to guide us with what directions we

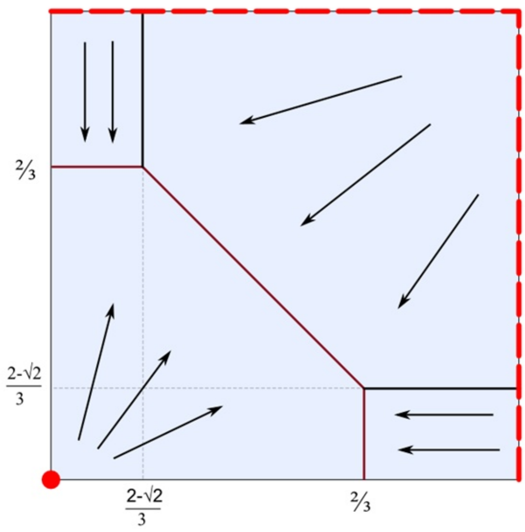

Fig. 6. The directions of measure transport used in our coupling γ for Example 5.

should use to transport measure in our coupling, in each region separately, as determined by the gradient of the utility function induced by the conjectured optimal mechanism.

In fact, we can reverse-engineer these steps when we do not have a conjecture about the optimal mechanism. Let us go back to Example 4, where we had two independent Beta items with parameters $(3, 3)$ and $(3, 4)$. Also, recall Figure 3 where we show the measure μ induced by the product of these distributions, according to (10).

How might the optimal mechanism for this example look like? It is reasonable to try to find a mechanism with the following properties:

—For sufficiently small values of v_1 and sufficiently large values of v_2 the mechanism offers item 2 with probability 1 and item 1 with probability strictly smaller than 1. Let us denote $\mathcal{A}$ the unknown subset of the typeset where this may happen.

—For sufficiently small values of v_2 and sufficiently large values of v_1 the mechanism offers item 1 with probability 1 and item 2 with probability strictly smaller than 1. Let us denote $\mathcal{B}$ the unknown subset of the typeset where this may happen.

Now let us revisit Condition (16). If we were to use this condition to show optimality of our yet-unknown mechanism, we would certainly be allowed to:

(1) transport measure southward in region $\mathcal{A}$; and

(2) transport measure westward in region $\mathcal{B}$.

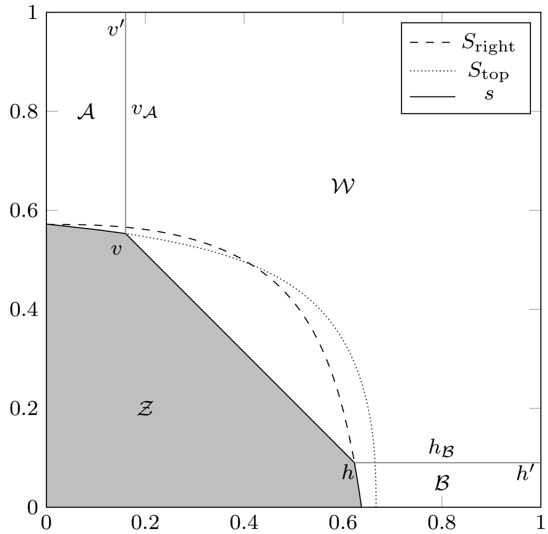

Fig. 7. Reverse-engineering the optimal mechanism from μ.

We may also be able to push measure eastward in region $\mathcal{A}$, if item 1 is allocated with probability 0 in this region. Similarly, we may be able to push measure northward in region $\mathcal{B}$, if item 2 is allocated with probability 0 in this region. As we want to be versatile, let us ignore these extra possibilities, insisting on southward transports in region $\mathcal{A}$ and westward transports in region $\mathcal{B}$.

With these restrictions on our transports let revisit measure μ, trying to close in on subsets of X that regions $\mathcal{A}$ and $\mathcal{B}$ may occupy. In Figure 7, we have drawn a monotone strictly concave curve S_{top} that traces, for each x_1, the top-most point (x_1, x_2) such that, restricted to the vertical segment between points (x_1, x_2) and $(x_1, 1)$, $\mu_+ \succeq_1 \mu_-$, where $\succeq_1$ denotes first-order stochastic dominance between measures defined as follows.

Definition 2. *If μ_1, μ_2 are two non-negative measures defined on some $S \subseteq X$ such that $\mu_1(S) = \mu_2(S)$, we say that μ_1* first-order stochastically dominates *μ_2, denoted $\mu_1 \succeq_1 \mu_2$, iff there exists a coupling $\gamma \in \Gamma(\mu_1, \mu_2)$ between μ_1 and μ_2 such that, almost everywhere with respect to $\gamma(x, y)$, x is coordinate-wise larger than or equal to y. Equivalently, for all non-decreasing measurable functions u, $\int_S u d\mu_1 \geq \int_S u d\mu_2$.*

Given that $X_+ \setminus \{\vec{0}\}$ is an increasing set, point $S_{\text{top}}(x_1)$ must belong to X_-, if it exists, and it can be identified by finding the largest x_2 such that the total measure on the segment between points (x_1, x_2) and $(x_1, 1)$ under μ is 0. Notice that for some x_1's there fails to be a segment with these properties, so S_{top} is undefined for those x_1's. Similarly, the monotone strictly concave curve S_{right} traces, for each x_2, the rightmost point (x_1, x_2) such that, restricted on the horizontal segment between points (x_1, x_2) and $(1, x_2)$, $\mu_+ \succeq_1 \mu_-$. Again, this curve is not defined for all x_2, and it lies within X_-.

To accommodate Constraints (1) and (2) on our transports of measure in the (still unidentified) regions $\mathcal{A}$ and $\mathcal{B}$, it suffices to pick $\mathcal{A}$ to be any region defined by $S_{\rm top}$, the left and top boundaries of $[0,1]^2$ and a vertical segment $v_{\mathcal{A}} = (vv')$ for some $v \in S_{\rm top}$ and v' on the top edge of $[0,1]^2$, as in the figure. Similarly, we can pick $\mathcal{B}$ to be any region defined by $S_{\rm right}$, the bottom and right boundaries of $[0,1]^2$ and a horizontal segment $h_{\mathcal{B}} = (hh')$, for some $h \in S_{\rm right}$ and h' on the right edge of $[0,1]^2$. For any region $\mathcal{A}$ as above, by the definition of $S_{\rm top}$, we can couple μ_+ restricted in $\mathcal{A}$ with μ_- restricted in $\mathcal{A}$ while respecting Constraint (1), as we can do this separately for every "vertical slice" of $\mathcal{A}$. Similarly for any $\mathcal{B}$ as above, we can couple μ_+ restricted in $\mathcal{B}$ with μ_- restricted in $\mathcal{B}$ while respecting Constraint (2).

The next question is how to pick segments $v_{\mathcal{A}}$ and $h_{\mathcal{B}}$, and what to do with the rest, $X \setminus \mathcal{A} \cup \mathcal{B}$, of the typeset. A natural approach is to assume that the remainder of the typeset is partitioned into two regions, $\mathcal{W}$ and $\mathcal{Z}$, of types that will purchase the bundle $\{1,2\}$ of both items at some price $p_{\{1,2\}}$ and of types that will not purchase anything. Clearly, the utility in $\mathcal{W}$ will be of the form $u(x) = x_1 + x_2 - p_{\{1,2\}}$ while the utility in $\mathcal{Z}$ will be $u(x) = 0$. Hence, the boundary between these regions must be a 45° segment.

Given this, a natural way to finish would be to try to identify a point $v \in S_{\rm top}$ and $h \in S_{\rm right}$ such that:

(i) The straight segment, (vh), between points v and h has a 45° angle; more precisely, we want that $(h_1 - v_1, h_2 - v_2) \propto (1, -1)$.

(ii) The region, $\mathcal{Z}$, under the curve S defined by the initial part of $S_{\rm top}$ (between points $S_{\rm top}(0)$ and v), the straight segment (vh), and the initial part of $S_{\rm right}$ (between points h and $S_{\rm right}(0)$) has total measure $\mu(\mathcal{Z}) = 0$ and is convex;

(iii) The region, $\mathcal{W}$, enclosed by the straight segments $v_{\mathcal{A}}$, $(v'\vec{1})$, $(\vec{1}h')$, $h_{\mathcal{B}}$ and (vh), satisfies $\mu_+|_{\mathcal{W}} \succeq_1 \mu_-|_{\mathcal{W}}$, where $\mu_+|_{\mathcal{W}}$ and $\mu_-|_{\mathcal{W}}$ denote respectively the restrictions of measures μ_+ and μ_- in region $\mathcal{W}$.

Note that, if we can identify points v and h satisfying Requirements (i)–(iii), we are done. Indeed, let us define the function $u : x \mapsto \ell_1(x, \mathcal{Z})$, mapping each type x to its ℓ_1 distance from set $\mathcal{Z}$. Clearly:

—From (ii), $\mathcal{Z}$ is convex, hence $u(x)$ is also convex. Given that $\mathcal{Z}$ is a decreasing set, u is non-decreasing. It is also clearly non-negative and 1-Lipschitz. So u is feasible for (P). If we could also find a coupling γ between μ_+ and μ_- that satisfies Condition (16), this would establish that u is optimal for (P). We proceed to do this next separately for μ_+ and μ_- restricted to each region.

—**Region $\mathcal{Z}$:** $u(x) = 0$, for all $x \in \mathcal{Z}$, conforming to our intention that $\mathcal{Z}$ is the set of types that do not purchase anything. Given that $\nabla u(x) = 0$, for all $x \in \mathcal{Z}$, Condition (16) implies that we are allowed to transport measure in north-east directions in this region. Given that $\mathcal{Z}$ lies entirely within $X_- \cup \{\vec{0}\}$, μ_+ only resides at $(0,0)$. Moreover, $\mu(\mathcal{Z}) = 0$. Hence, it is possible to couple μ_+ and μ_- in $\mathcal{Z}$ with only north-east transports.

—**Region $\mathcal{W}$:** $u(x) = x_1 + x_2 - p^*$, for all $x \in \mathcal{W}$, where p^* is the intercept of segment (vh) if it were extended to hit the x_1 axis. This conforms to our intention that the types in $\mathcal{W}$ purchase the grand bundle. Given that $\nabla u(x) = 1$, for all

$x \in \mathcal{W}$, Condition (16) implies that we can transport measure in south-west directions in this region. Given that $\mu_+|_\mathcal{W} \succeq_1 \mu_-|_\mathcal{W}$, as per Requirement (iii), we can indeed couple μ_+ and μ_- in this region with only transports in south-west directions, by Definition 2.

—**Region $\mathcal{A}$:** Given that $\mathcal{Z}$ is convex (and therefore the lower boundary of region $\mathcal{A}$ is less steep that $45°$ as per the location of this boundary with respect to the segment (vh)), for all $x \in \mathcal{A}$, $u(x)$ is the vertical distance between x and S_{top}. So $\frac{\partial u}{\partial x_2} = 1$, which conforms to our intention that item 2 should be allocated with probability 1 in this region. Moreover, this means by Condition (16) that we are allowed to transport measure southward in our coupling between μ_+ and μ_- in this region. And, by the definition of S_{top}, restricted to this region μ_+ and μ_- can be coupled with only southward transports.

—**Region $\mathcal{B}$:** Similarly, given that $\mathcal{Z}$ is convex (and therefore the left boundary of region $\mathcal{B}$ is steeper that $45°$), $u(x)$ is the horizontal distance between x and S_{right} for all $x \in \mathcal{B}$. Via similar arguments as those employed for region $\mathcal{A}$, μ_+ and μ_- can be coupled in this region with only westward transports, respecting Condition (16).

By the above discussion, if we can identify points v and h satisfying Requirements (i)–(iii), this means that we have identified our optimal mechanism. It turns out that, for our specific distributions, Requirements (i)—(iii) can be satisfied and we can analytically compute the points v and h, as shown in Figure 7. So we have managed to reverse-engineer the optimal mechanism for our setting by exploiting Condition (16). The mechanism can be described indirectly by specifying $\mathcal{Z}$. In terms of $\mathcal{Z}$, the utility function of the optimal mechanism is $u : x \mapsto \ell_1(x, \mathcal{Z})$. From u we can also find the lotteries offered by the optimal mechanism: The types in $\mathcal{W}$ receive both items with probability 1. Each type $x \in \mathcal{A}$ receives item 2 with probability 1 and item 1 with probability that equals minus the slope of S_{top} at point $S_{\text{top}}(x_1)$. Similarly, each type $x \in \mathcal{B}$ receives item 1 with probability 1 and item 2 with probability that equals minus the inverse slope of S_{right} at point $S_{\text{right}}(x_2)$.

Some remarks are in order before we conclude this section:

—First, as a byproduct of our derivation above we have proven our claim in Example 4 that the optimal mechanism offers an uncountably large menu of lotteries. Indeed, recall that S_{top} is strictly concave. Hence, the allocation probability of item 1 differs in every vertical strip within this region. Thus, a continuum of lotteries are offered to the types in $\mathcal{A}$. The same is true for the types in $\mathcal{B}$.

—On the other hand, as promised in Section 3.3, there does exist a succinct description of the optimal mechanism. All we need to maintain is an analytic description of the boundary of region $\mathcal{Z}$.

—We emphasize again that the approach followed in this section to reverse-engineer the optimal mechanism is not guaranteed to succeed. Indeed, it is based on a weak dual (D') of our optimal mechanism design formulation (P). Based on this weak dual, it identifies a complementary slackness condition, (16), which ties solutions to (P) and (D') in a particular way. It then makes guesses about the optimal mechanism, and tries to follow through with these guesses using Condition (16). Despite the fact that it is not guaranteed to succeed, the approach is quite successful in identifying optimal mechanisms. In [Tzamos 2015], Christos Tzamos provides an applet where the heuristic approach of this section is applied to reverse-engineer the optimal mechanism for user-specified Beta distributions.

—Given that the heuristic method proposed in this section may not succeed, it remains important to develop a technique that is guaranteed to succeed. This is what we do in the next section.

5.7 A Strong Duality Theorem

In previous sections, we identified a weak dual (D') of our formulation, (P), of optimal mechanism design as a convex optimization problem. Based on the weak dual, in Section 5.6.1, we gave an example where we were able to certify the optimality of a mechanism. In Section 5.6.2, we also showed how we can use the weak dual to reverse-engineer the optimal mechanism. Still, (D') is a weak dual, so there are settings where both certifying and reverse-engineering the optimal mechanism will fail [Daskalakis et al. 2015].

We would thus like to obtain a strong dual of formulation (P), whose optimum is guaranteed to match that of (P). The advantage of a strong dual would be that, for every mechanism design problem (n, F), the optimal mechanism u for (P) would have a certificate of optimality in the form of a solution to the dual. Thus, we would know what type of certificate to look for, and we would also be able to reverse-engineer optimal mechanisms by exploiting the structure of those certificates.

The challenge we encounter though is that a priori it is not clear if a tight dual to (P) should exist. Indeed, (P) is an infinite-dimensional optimization problem and contains constraints on the variable u, such as convexity, which are non-standard. In particular, we are not aware of a duality framework, based on infinite-dimensional linear programming, that can accommodate such constraints directly. We prove our own extension of Monge-Kantorovich duality to accommodate the constraints of (P), establishing the following.

THEOREM 4 [DASKALAKIS ET AL. 2015]. *Formulation (P) has a strong dual formulation (D), taking the form of an optimal transport problem, as follows:*

$$(P): \sup_{\substack{u \text{ satisfies } (11), \\ (12) \text{ and } (13)}} \int_X u d\mu = \inf_{\substack{\gamma \in \Gamma(\mu'_+, \mu'_-) \\ \mu'_+ - \mu'_- = \mu' \succeq_{cvx} \mu}} \int_{X \times X} |(x-y)_+|_1 d\gamma : (D)$$

Moreover, both sup *and* inf *are attainable.*

We next explain the statement of Theorem 4, comparing (D) to the weak-dual (D') of Section 5.5. We then discuss the point of deriving a strong dual, before turning to some important applications of our strong duality theorem.

5.7.1 *Discussion of Theorem 4.* First, to clarify the notation used in Theorem 4, $\succeq_{cvx}$ is the standard notion of convex dominance between measures, defined as follows.

DEFINITION 3. *For two measures* μ_1 *and* μ_2 *over* X *we say that* μ_1 convexly dominates μ_2*, denoted* $\mu_1 \succeq_{cvx} \mu_2$*, iff, for all measurable, non-decreasing and convex functions* u*,* $\int_X ud\mu_1 \geq \int_X ud\mu_2$.

Compared to the first-order stochastic dominance of Definition 2, convex dominance is a weaker requirement as it only requires that $\int_X ud\mu_1 \geq \int_X ud\mu_2$ for convex, non-decreasing u's, as opposed to all non-decreasing u's.

As non-decreasing, convex functions model utility functions of risk-seeking individuals, when μ_1, μ_2 are probability measures and $\mu_1 \succeq_{cvx} \mu_2$, this means that any risk-seeking individual prefers a prize distributed according to μ_1 to a prize distributed according to μ_2. More generally, when $\mu_1 \succeq_{cvx} \mu_2$, this essentially means that μ_1 can be obtained from μ_2 through a sequence of operations allowed to:

—send (positive) measure to coordinate-wise larger points—this makes the integral $\int ud\mu_1$ larger than $\int ud\mu_2$ since u is non-decreasing; or

—spread (positive) measure so that the mean is preserved—this makes the integral $\int ud\mu_1$ larger than $\int ud\mu_2$ since u is convex.

Now that our understanding of convex dominance is grounded, let us compare our strong dual (D) to the weak-dual (D') of Section 5.5. The two problems are very similar. They are both min-cost transportation problems, seeking a coupling $\gamma(x, y)$ between two non-negative measures under which the total cost under the same cost function $|(x - y)_+|_1$ is minimized. The difference between the two problems is that (D') seeks a coupling between μ_+ and μ_-, the positive and negative parts of measure μ derived from F according to (10). (D) is also allowed to pre-process μ without incurring any cost before seeking a coupling between the positive and negative parts of the processed measure. In particular, (D) is allowed to choose any measure $\mu' \succeq_{cvx} \mu$ and couple the positive and negative parts of that measure μ'. This way it may reduce the transportation cost, which Theorem 4 says is guaranteed to match the optimum of (P).

5.7.2 *The Point of Theorem 4.* So, what is the point of obtaining a strong dual of our mechanism design problem? We have already alluded to the benefits of strong duality above. By analogy, these should also be clear to anyone familiar with the uses of strong linear programming duality in combinatorial optimization. Let us expand a bit. In comparison to our weak dual (D'), our strong dual (D) is more powerful as it allows us to certify the optimality of *any* mechanism. In particular, for any mechanism design problem defined by some n and F, we are guaranteed to find a measure $\mu' \succeq_{cvx} \mu$ and a coupling γ between μ'_+ and μ'_- whose transportation cost equals the optimal revenue. In particular, we can identify conditions tying a solution u to (P) with a solution (μ', γ) to (D) that, whenever satisfied, establish the joint optimality of both u and (μ', γ). Additionally, using these conditions as a guiding principle, we can reverse-engineer optimal mechanisms in settings where we do not have a conjecture about what the optimal mechanism is, as we did in Section 5.6.2 using weak duality. Except now this approach is guaranteed to work. We refer the interested reader to [Daskalakis et al. 2015] for these conditions, as well as examples where they are used to certify optimality of mechanisms. The approach is similar to Sections 5.6.1 and 5.6.2, so we do not expand more here.

We prefer to turn instead to exciting applications of the strong duality theorem to obtain characterization results.

5.8 Characterizing Optimal Mechanisms

Using our strong duality theorem (Theorem 4) to certify optimality of mechanisms may be cumbersome. A pertinent question is this: Can we somehow use the theorem behind the scenes to develop simple characterization results of mechanism optimality? In this section, we obtain such a characterization: given a proposed mechanism for some setting of n and F, we identify a collection of *necessary and sufficient* conditions on μ for the mechanism to be optimal. In particular, these conditions do not involve the dual problem at all. They are just conditions on the measure μ derived from n and F using (10). We proceed to give a flavor of our characterization, and some applications. We start with a characterization of grand-bundling optimality, proceeding to general mechanisms afterwards.

5.8.1 *Characterization of Grand-Bundling Optimality.* Let us complement Definition 3 with the following definition.

Definition 4. *For two measures μ_1 and μ_2 over X we say that μ_1* second-order stochastically dominates *μ_2, denoted $\mu_1 \succeq_2 \mu_2$, iff for all measurable, non-decreasing and concave functions u, $\int_X u d\mu_1 \geq \int_X u d\mu_2$.*

As concave and non-decreasing functions model utility functions of risk-averse individuals, if μ_1, μ_2 are probability measures and $\mu_1 \succeq_2 \mu_2$, this means that any risk-averse individual prefers a prize distributed according to μ_1 to a prize distributed according to μ_2. More generally, $\mu_1 \succeq_2 \mu_2$ essentially means that μ_1 can be obtained from μ_2 via a sequence of operations that shift positive measure to coordinate-wise smaller points and do mean-preserving merges of positive measure.

With the above definition, we can use our duality theorem behind the scenes to obtain the following characterization of grand-bundling optimality.

Theorem 5 [Daskalakis et al. 2015]. *For all n and F, the mechanism that only offers the bundle of all items at some price p is optimal if and only if measure μ defined by (10) satisfies $\mu|_{\mathcal{W}} \succeq_2 0 \succeq_{cvx} \mu|_{\mathcal{Z}}$, where $\mathcal{W}$ is the subset of types that can afford the grand bundle at price p, $\mathcal{Z}$ the subset of types who cannot, and $\mu|_{\mathcal{W}}$, $\mu|_{\mathcal{Z}}$ are respectively the restrictions of μ in subsets $\mathcal{W}$ and $\mathcal{Z}$.*

Sufficient conditions for grand-bundling optimality have been an active line of inquiry—see e.g. [Manelli and Vincent 2006; Pavlov 2011]. Theorem 5 provides a single condition, in the form of two stochastic dominance relations, that is necessary and sufficient.

As a corollary of this theorem, we can show Theorem 1 of Section 3.1, pertaining to the counter-intuitive behavior of the optimal mechanism for n i.i.d., uniform $[c, c+1]$ items:

- The first part of this theorem is an extension of Pavlov's result for $n = 2$ [Pavlov 2011]. Its proof is by a geometric construction showing that the sufficient condition will be met for all n's as long as c is large enough.
- The second, and more counter-intuitive, part of the theorem is shown by arguing that all prices p will fail to satisfy $0 \succeq_{cvx} \mu|_{\mathcal{Z}}$. Essentially, what happens in this

setting is that all the positive measure is surface measure on the facets $x_i = c+1$, and there is an atom of $+1$ at the origin $\vec{c}$. All other facets and the interior of the cube $[c, c+1]^n$ have negative measure. Given this, if n is large enough, measure $\mu|_{\mathcal{Z}}$ exhibits a phase transition in terms of the price p: if p is large enough that $\mathcal{Z}$ has a positive-measure intersection with facet $x_1 = c+1$, function $u(x) = \mathbf{1}_{\{x_1=c+1\}}$ is an explicit witness that constraint $0 \succeq_{cvx} \mu|_{\mathcal{Z}}$ is violated, as $\int u d\mu|_{\mathcal{Z}} > 0$. On the other hand, if p is small enough that $\mathcal{Z}$ has measure 0 intersection with this facet, there turns out not to be enough negative mass in region $\mathcal{Z}$ to balance the positive atom at $\vec{c}$, as required for $0 \succeq_{cvx} \mu|_{\mathcal{Z}}$ to hold.

5.8.2 *Characterization of General Mechanisms.* Our characterization of grand-bundling optimality from the previous section naturally extends to arbitrary mechanisms. Let us briefly discuss this generalization, referring the reader to [Daskalakis et al. 2015] for more details.

Consider a mechanism $\mathcal{M}$ for some setting n and F. $\mathcal{M}$ induces a partition of the typeset X into subsets of types that will decide to purchase different lotteries in the menu offered by $\mathcal{M}$. Assuming that $\mathcal{M}$ offers a finite number of lotteries, this partition may look like Figure 8, where each cell c corresponds to a subset of types that will purchase the same lottery (p^c, t^c), where p^c is a vector of allocation probabilities and t^c a price.

Fig. 8. A partition of the typeset induced by some finite menu of lotteries.

If the mechanism is a grand-bundling mechanism, then there are only two regions and Theorem 5 defines a pair of stochastic dominance conditions that are necessary and sufficient for its optimality. Naturally, our generalization to general mechanism will require one condition per cell c of the partition. To describe them, let us define vector $\vec{v}^c$ in terms of p^c as follows:

$$\forall i : v_i^c = \begin{cases} 1, & \text{if } p_i^c = 0 \\ -1, & \text{if } p_i^c = 1 \\ 0, & \text{otherwise} \end{cases}$$

Moreover, for a vector $\vec{v} \in \{-1, 1, 0\}^n$, let us define a stochastic dominance relation with respect to $\vec{v}$ as follows:

Definition 5. *For a vector $\vec{v} \in \{-1, 1, 0\}^n$ and two measures μ_1 and μ_2 over X, we say that μ_1* convexly dominates *μ_2* with respect to *$\vec{v}$, denoted $\mu_1 \succeq_{cvx(\vec{v})} \mu_2$, iff, for all measurable and convex functions u that are non-decreasing in all coordinates i such that $v_i = 1$ and non-increasing in all coordinates i such that $v_i = -1$: $\int_X u d\mu_1 \geq \int_X u d\mu_2$.*

Clearly, $\mu_1 \succeq_{cvx(\vec{1})} \mu_2 \Leftrightarrow \mu_1 \succeq_{cvx} \mu_2$ and it easily follows that $\mu_1 \succeq_{cvx(-\vec{1})} \mu_2 \Leftrightarrow \mu_2 \succeq_2 \mu_1$. So Theorem 5 can be restated as specifying the following necessary and sufficient conditions for grand-bundling optimality: $0 \succeq_{cvx(\vec{1})} \mu|_{\mathcal{Z}}$ and $0 \succeq_{cvx(-\vec{1})} \mu|_{\mathcal{W}}$, where $\mathcal{Z}$ and $\mathcal{W}$ are the subsets of types that purchase nothing and the grand bundle respectively.

Our general characterization is the following:

Characterization of General Mechanisms

$$(\mathcal{M} \text{ is optimal}) \Longleftrightarrow \begin{pmatrix} \text{for every cell } c \text{ in the partition of } X \text{ in-} \\ \text{duced by } \mathcal{M}\text{: } 0 \succeq_{cvx(\vec{v}^c)} \mu|_c, \text{ where } \mu|_c \\ \text{is the restriction of } \mu, \text{ defined in (10), to} \\ \text{cell } c \end{pmatrix}.$$

The interested reader is referred to [Daskalakis et al. 2015] for more details.

5.9 Concluding Remarks

In Sections 5.1–5.8 we took up the challenge of understanding the structure of multi-item mechanisms, trying to demystify their counter-intuitive behavior identified in Section 3.

Our contribution was a duality framework based on which we can certify the optimality of mechanisms in every single-buyer setting. In particular, we showed that the optimal mechanism design problem has a tight dual problem, taking the form of an optimal transportation problem. Given a proposed mechanism, we can test if it is optimal by identifying solutions to the dual achieving the same value, as we did in Section 5.6.1, except that more generally we can also pre-process μ with cost-free convex shuffling operations before finding our transports, as allowed by the strong dual (D). Given our strong duality, dual certificates are guaranteed to exist. Moreover, if we do not have a conjectured optimal mechanism for some setting of interest, we can exploit the dual to reverse-engineer it, as we did in Section 5.6.2. Finally, we developed "duality-theory oblivious" tools, characterizing the optimality of mechanisms in terms of the buyer's distribution. These tools use the power of the duality framework behind the scenes, presenting clean necessary and sufficient conditions for a mechanism to be optimal.

All in all, we believe that our framework provides a new perspective on multi-item mechanisms, presenting a tool whose applicability is universal, given our strong duality. Our work opens interesting lines for future investigation, and we are particularly interested in:

(1) extending the duality framework to accommodate multiple buyers; and

(2) developing technical machinery to facilitate testing stochastic dominance relations.

While we recognize that a lot remains to be done, we believe that the afore-described framework represents the beginning of a principled approach towards a structural understanding of multi-item multi-buyer mechanisms.

6. THE COMPUTATIONAL COMPLEXITY OF MULTI-ITEM MECHANISMS

6.1 Philosophy

The duality based framework of the previous section targeted closed-form characterizations of optimal mechanisms. A related goal is to study the computational complexity of optimal mechanisms. In particular, we are interested in whether optimal mechanisms can be computed and implemented computationally efficiently. There are several reasons why this question is important:

—First, if optimal mechanisms were computationally intractable, then why should practitioners care about them, especially in settings with a large number of bidders or items? Computational intractability would justify using approximate mechanisms in practical applications.

—Moreover, as we have seen, getting closed-form descriptions of optimal mechanisms is a challenging task. Despite intense work in the literature, including that of the previous section, we are still far from characterizing optimal multi-bidder mechanisms. When closed-form descriptions are unknown, being able to compute optimal mechanisms is an interesting middle ground.

—Additionally, computing optimal mechanisms would be a great tool for researchers who want to gain familiarity with the structure of these mechanisms and/or test hypotheses about their structure or performance of approximate solutions.

—Finally, one might expect that studying optimal mechanism design from an algorithmic point of view may reveal structure that might be hard to observe using non-algorithmic tools.

For all the above reasons, we find it important to study the computational complexity of optimal mechanisms. Over the past few years, our and other groups have made tremendous progress in the complexity of optimal mechanism design [Cai and Daskalakis 2011; Cai et al. 2012a; Alaei et al. 2012; Cai et al. 2012b; Daskalakis et al. 2012; Cai and Huang 2013; Cai et al. 2013a; 2013b; Alaei et al. 2013; Bhalgat et al. 2013; Daskalakis et al. 2014; Chen et al. 2014; Daskalakis and Weinberg 2015; Daskalakis et al. 2015]. See also [Hartline 2013; Chawla and Sivan 2014; Cai et al. 2015] for recent surveys, covering some of this work. We proceed to give a flavor of what is known, restricting our attention to multi-item multi-bidder settings with additive bidders. As discussed below, all our results in this section extend to much broader settings.

6.2 Setting

In this section, we restrict our attention to a seller with n items and m additive bidders interested in those items. The type of each bidder is a vector t_i of values for the items, which are jointly distributed according to some distribution F_i. In particular, t_{ij} will denote the value of bidder i for item j. We assume that the distribution F_i is known to the seller and all the other bidders, but only bidder i

knows his realized type. We also assume that bidder types are independent and use $\vec{t}$ to denote the types of all bidders, also called the *type profile*. Finally, given a type vector $\vec{t}$ and i, we denote, as is customary, by $\vec{t}_{-i}$ the vector containing the types of all bidders except the type of bidder i. Accordingly, we use $(t_i; \vec{t}_{-i})$ to denote $\vec{t}$.

The goal of the seller is to design a mechanism that optimizes his expected revenue, when the expectation is computed with respect to the types of the bidders, the randomness in the mechanism, if any, and the randomness in the strategies of the bidders, if any. The mechanism is allowed to be any protocol that interacts with the bidders and has the bidders interact with each other in some fashion. Whatever the protocol is, it is supposed to output an allocation $x \in \{0,1\}^{mn}$ of items to bidders, where x_{ij} indicates whether item j is given to bidder i. As we assume that there is exactly one copy of each item, any allocation ever output by the mechanism must satisfy that

$$\sum_i x_{ij} \leq 1, \text{for all items } j.$$

We call $\mathcal{F}$ the subset of $\{0,1\}^{mn}$ satisfying the above constraints. The mechanism can also charge prices, as long as the bidders accept to pay those prices.

While it is hard to optimize over protocols, it follows from the revelation principle that it suffices to optimize over a simpler class of mechanisms called "direct mechanisms." These mechanisms are described by an allocation function $X : \vec{t} \mapsto \Delta(\mathcal{F})$, mapping type profiles to distributions over feasible allocations, and a price function $P : \vec{t} \mapsto \Delta(\mathbb{R}^m)$ mapping type profiles to distributions over price vectors, and are implemented as follows:

—The bidders are asked to report their types to the mechanism.

—If $\vec{t}$ are the reported types, the mechanism samples $x \sim X(\vec{t})$ and $p \sim P(\vec{t})$, implements allocation x and charges prices according to p.

For convenience, we will use $\mathcal{X}(\vec{t})$ to denote a random variable distributed according to $X(\vec{t})$, and similarly $\mathcal{P}(\vec{t})$ to denote a random variable distributed according to $P(\vec{t})$.

While, in principle, the bidders need not be truthful about their types, due to the revelation principle we can also assume without loss of generality that it will be in their best interest to do so. We may also assume that it is not hurtful to them to participate in the mechanism. In particular, we may assume that the mechanism is *Bayesian Incentive Compatible* and satisfies *Individual Rationality* according to the following definition.

Definition 6. *We say that a direct mechanism (X, P) is* Bayesian Incentive Compatible (BIC) *iff for all bidders i and types t_i and t'_i in the support of F_i:*

$$\mathbb{E}_{\vec{t}_{-i}}[t_i \cdot \mathcal{X}(\vec{t}) - \mathcal{P}(\vec{t})] \geq \mathbb{E}_{\vec{t}_{-i}}[t_i \cdot \mathcal{X}(t'_i; \vec{t}_{-i}) - \mathcal{P}(t'_i, \vec{t}_{-i})]. \tag{17}$$

We say that a direct mechanism (X, P) satisfies Individual Rationality *or* is IR *iff for all bidders i and types t_i in the support of F_i:*

$$\mathbb{E}_{\vec{t}_{-i}}[t_i \cdot \mathcal{X}(\vec{t}) - \mathcal{P}(\vec{t})] \geq 0. \tag{18}$$

(17) expresses that the expected utility of bidder i cannot be improved by misreporting his type, while (18) expresses that the expected utility of bidder i is non-negative if he is truthful. Both expectations are with respect to the types of the other bidders as well as the randomness in the allocation and price functions, if any.

Under the assumption that bidders will report their types truthfully to a BIC, IR mechanism, its revenue can be expressed as follows:

$$\mathbb{E}_{\vec{t}}\left[\sum_i \mathcal{P}_i(\vec{t})\right]. \tag{19}$$

With all the context provided above, the seller is given distributions $F_1, \ldots, F_m$ over bidder types and seeks to compute a BIC, IR mechanism of optimal revenue.

6.3 Computationally Efficient in What Exactly?

When it comes to studying multi-item mechanisms from a computational perspective what we need to answer first is what the input is, and how its description complexity is measured. There are several ways one can go about this, including the following.

—One approach is to assume that the bidders' distributions come from a parametric family of distributions, and specify the parameters of each bidder's distribution. This approach would allow us to accommodate both discrete and continuous distributions. In this case, the description complexity of the distributions is the description complexity of all parameters required to describe them.

—Another approach is to assume that the distributions are discrete and provide them explicitly, by listing the elements in their support and the probability that each element is sampled. The explicit description is reasonable for distributions with a small and discrete support. Here the description complexity of the distributions is the description complexity of all the elements in their support and their associated probabilities.

—Finally, one can assume to have sample access to the distribution F_i of each bidder. Here each distribution can be thought of as a subroutine that a seller can call to get an independent sample from F_i. The description complexity of the distributions is harder to define in this model. One way to do this is to assume that the subroutines only output numbers of certain bit complexity, or truncate all numbers output by these subroutines to some accuracy. We do not want to dwell on this point too much though.

We will restrict our attention to the simplest model, assuming that all our distributions are explicit. The parametric and sample-access models are important to study as well, but we are not aware of efficiently computable mechanisms that can accommodate these models in reasonable generality without losing revenue. In fact, we should not expect to get general, exactly optimal algorithms in these settings given the following result:

THEOREM 6 [DASKALAKIS ET AL. 2014]. *Consider the algorithmic problem of designing an optimal mechanism for selling n items to a single, additive buyer whose values for the items are independently distributed, according to distributions* *that are supported on two rational numbers with rational probabilities. In particular, the distribution of each item i is specified by a pair of rational numbers $\{a_i, b_i\}$ and a rational probability p_i.*

Unless $\mathtt{ZPP} \supseteq \mathtt{P}^{\#\mathtt{P}}$*, any mechanism (of any type, direct or indirect) that can be computed in expected polynomial time cannot be both optimal and executable in expected polynomial time.*

Notice that the mechanism design problem in the statement of Theorem 6 conforms to the parametric model for describing distributions. And the theorem provides a particularly simple setting where, unless we spent super-polynomial time, we will not be able to compute an efficiently executable mechanism that is also optimal. As the sample-access model is not easier than the parametric one, the same result applies to this model as well. Finally, we will only remark here that the assumption $\mathtt{ZPP} \supseteq \mathtt{P}^{\#\mathtt{P}}$ is a standard complexity-theoretic assumption about the relations of complexity classes $\mathtt{P}$, $\mathtt{ZPP}$ and $\#\mathtt{P}$. The interested reader is referred to standard complexity theory textbooks for more discussion.

While Theorem 6 is quite discouraging, especially given the simplicity of the mechanism design problem that is shown intractable, it still does not preclude efficiently computable mechanisms that are near-optimal. In particular, it is an interesting open problem to determine whether there exist efficiently computable mechanisms that achieve a $(1 - \epsilon)$-fraction of the optimal revenue, for any desired accuracy $\epsilon > 0$, as long as one is willing to invest time polynomial in $1/\epsilon$ and the parameters of the distribution for their computation. This problem is open even in the simple setting of Theorem 6.

6.4 Computing Optimal Mechanisms, and a Bonus

We turn to the explicit model of representing bidder distributions, and ask whether optimal mechanisms can be computed efficiently. There is still a wrinkle we need to overcome however. As we said a direct mechanism is a pair of functions $X : \vec{t} \mapsto \Delta(\mathcal{F})$ and $P : \vec{t} \mapsto \Delta(\mathbb{R}^m)$. So to describe these functions explicitly, we need to specify a distribution over allocations and a distribution over price vectors for every possible type profile. This is problematic though as revealed by a little calculation. Suppose that the support of every bidder's type distribution has size k. Then to specify each distribution we need to give k numbers and $k - 1$ probabilities. So to describe all these distributions only requires $O(mk)$ numbers. On the other hand, there are k^m possible type profiles. So we cannot hope to compute X and P explicitly. Besides, even the outputs of these functions are distributions over high-dimensional spaces.

We thus need to be smart about how we represent mechanisms. We need to represent them implicitly, whilst still being able to compute on the implicit representation. In the additive setting that we consider here, it turns out that the so-called "reduced-form" is a good representation:

DEFINITION 7. *The* reduced form $(\hat{x}, \hat{p})$ *of a mechanism* (X, P) *is a collection of single-variate functions* $\hat{x}_i : T_i \to [0,1]^n, i = 1, \ldots, m,$ *and* $\hat{p}_i : T_i \to \mathbb{R}, i = 1, \ldots, m,$ *where T_i is the support of F_i, which are related to X and P as follows. For all i and types $t_i \in T_i$:*

—$\hat{x}_{ij}(t_i) = \mathbb{E}_{\vec{t}_{-i}}[\mathcal{X}_{ij}(t_i;\vec{t}_{-i})]$; *in particular, the j-th coordinate of $\hat{x}_i(t_i)$ is the probability that the mechanism allocates item j to bidder i, conditioning on his reported type being t_i and assuming that the other bidders report their types truthfully.*

—$\hat{p}_i(t_i) = \mathbb{E}_{\vec{t}_{-i}}[\mathcal{P}_i(t_i;\vec{t}_{-i})]$; *in particular, $\hat{p}_i(t_i)$ is the expected price charged to bidder i, conditioning on his reported type being t_i and assuming that the other bidders report their types truthfully.*

The reduced form is called "reduced" because it is losing information about the mechanism. It can be viewed as projecting (X, P), which is a high-dimensional object, to a lower-dimensional space. So maintaining mechanisms in their reduced forms, creates two computational challenges:

(1) Given a reduced form $(\hat{x}, \hat{p})$, is it possible to verify computationally efficiently whether it is feasible, that is whether there exists an actual mechanism (X, P) whose reduced form agrees with $(\hat{x}, \hat{p})$?

(2) Given a reduced form $(\hat{x}, \hat{p})$ that is feasible, is it possible to implement some mechanism with this reduced form computationally efficiently?

Border and Che et al. have provided a collection of linear constraints that are necessary and sufficient for reduced-form feasibility [Border 1991; 2007; Che et al. 2011]. See also [Hart and Reny 2014]. These constraints have a nice interpretation as max-flow/min-cut constraints in a related flow network. However, they cannot be used towards computationally efficient algorithms for the above problems, as they are exponentially many. We can improve these results as follows:

Theorem 7 [Cai et al. 2012a; Alaei et al. 2012]. *The answers to both questions above are "yes." Namely, given a reduced form $(\hat{x}, \hat{p})$, we can verify in polynomial time whether it is feasible. Moreover, given a feasible reduced form $(\hat{x}, \hat{p})$, we can compute in polynomial time a pair of polynomial-time algorithms for sampling the allocation and price functions of a mechanism (X, P) whose reduced form is $(\hat{x}, \hat{p})$.*

Theorem 7 is very handy as it allows us to formulate polynomial-size linear programs for finding optimal mechanisms. In particular, it is not hard to see the following:

—The set of all possible reduced forms is convex, as they are projections of allocation and price functions, which themselves are a convex set as distributions over deterministic allocation and price functions.

—Given Theorem 7, there exists a polynomial-time algorithm that determines feasibility of reduced forms. This gives a computationally efficient separation oracle for the set of feasible reduced forms.

—The expected revenue of a mechanism can be expressed as a linear function of its reduced form.

—The BIC and IR constraints are also expressible as linear constraints in the reduced form.

—Finally, Theorem 7 implies that, given the reduced form of a mechanism, we can efficiently implement it.

Using the Ellipsoid algorithm, the above observations culminate to the following result:

Theorem 8 [Cai et al. 2012a]. *Consider a mechanism design setting with m additive bidders and n items. Given an explicit description of the bidders' type distributions, we can compute and implement an optimal mechanism in polynomial-time.*

Our new theorem provides an important counterpart to our structural results from Section 5, encompassing multi-bidder settings as well. In fact, the theorem can be generalized to much broader settings involving more complex allocation constraints [Cai et al. 2012b], budgets [Bhalgat et al. 2013; Daskalakis et al. 2015], non-additive bidders [Cai et al. 2013b], and even going beyond revenue and welfare to other interesting objectives, such as maximizing fairness or minimizing makespan [Daskalakis and Weinberg 2015]. Taken together our algorithmic results provide a very crisp understanding of the complexity of Bayesian mechanism design. In a recent article, we provide a concise overview of our work on this front, and refer the interested reader to this overview [Cai et al. 2015], as well as the surveys mentioned earlier [Hartline 2013; Chawla and Sivan 2014].

Let us conclude this section with a small treat. We promised earlier that the algorithmic perspective might actually reveal structure in the optimal mechanism that could be hard to see otherwise. While developing our algorithmic framework, we encountered remarkable structure in the optimal mechanism that is worth sharing here. Recall, from Fact 3 that Myerson's optimal single-item mechanism is a virtual welfare maximizer. It turns out that this holds in arbitrary settings. Here is the structure of the optimal mechanism for the setting of Theorem 8.

Theorem 9 [Cai et al. 2012b]. *When n items are sold to m additive bidders, the optimal mechanism is a* virtual welfare maximizer*, namely:*

(*1*) *The bidders are asked to report their types to the mechanism. Say that the reported types are $t_1, \ldots, t_m$.*
(*2*) *The reported types are transformed into virtual types, $h_1(t_1), \ldots, h_m(t_m)$, where each $h_i : T_i \to \mathbb{R}^n$ maps an additive type t_i to another additive type $h_i(t_i)$.*
(*3*) *Each item is allocated to the bidder with the highest virtual value for this item, with some lexicographic tie-breaking.*
(*4*) *Finally, prices are charged so that the mechanism is BIC.*

So the optimal mechanism has the exact same form in multi-item settings as it has in single-item settings! The differences between Theorem 9 and Fact 3 are the following:

—In Myerson's single-item setting, the virtual transformations h_i are deterministic, while in the multi-item setting they are actually randomized. Of course, we knew this had to be the case as by Fact 5 randomization is necessary in multi-item settings.
—As emphasized in Fact 3, in Myerson's setting, each h_i only depends on bidder i's distribution, but not on the other bidders' distributions and not even on how many other bidders show up. In the multi-item setting, each h_i may depend on

the distributions of all bidders, but importantly its argument is only bidder i's type.

Despite these differences, it is remarkable that the same structure applies to both single-item and multi-item settings. Again the above structure generalizes well-beyond the additive setting—see [Cai et al. 2013b; 2015].

7. CONCLUSIONS

In the past few decades, we witnessed the tremendous impact of Myerson's result [Myerson 1981] in mechanism design. Building on his result, we now understand virtually every aspect of revenue optimization in single-item settings. Moreover, we have seen numerous extensions, robustifying the result with respect to the details of the bidders' distributions, and expanding its applicability to a myriad of single-dimensional settings, accommodating budgets, online arrivals and departures of bidders, complex allocation constraints, non-linear designer objectives, and many more. Algorithmic and approximation techniques have played an important role in exploring these extensions, and indeed it has been thanks to the sharpness and simplicity of Myerson's result that this interplay between computation and mechanism design has been so fruitful. See [Hartline 2013; Chawla and Sivan 2014; Roughgarden 2015] for recent surveys of this literature.

Unfortunately, multi-item revenue optimization has not enjoyed the same fate due to our lack of understanding of multi-item mechanisms. Indeed, optimal multi-item mechanisms exhibit such rich structure that it is not clear whether there is a lens through which we can gain a crisp understanding of their properties. In this survey, we provided two approaches, based on duality theory and optimization, through which we obtained a fresh perspective on multi-item mechanism design. We have used these approaches to characterize the structure of multi-item mechanisms and showcased procedures, both analytical and algorithmic, via which the optimal mechanism can be identified.

We believe that the results presented in this survey have begun to resemble a cohesive theory of multi-item auctions, opening exciting directions for future investigation. To identify just a few:

- It is important to extend the duality-based framework of Section 5 to accommodate multiple bidders.
- What is the sensitivity of the structural and algorithmic results on the details of the bidder type distributions?
- In settings where the seller knows the bidder distributions, but the bidders do not, what is the Bayesian-optimal dominant strategy truthful mechanism?

Ultimately, we feel that the foundations have been laid for exciting developments in optimal multi-item mechanism design, expecting a lot more progress on this front in the next years.

Acknowledgments

We thank Yang Cai, Shaddin Dughmi, Christos Tzamos and Matt Weinberg for valuable comments on drafts of this survey. We thank Christos for all the figures.

REFERENCES

Alaei, S., Fu, H., Haghpanah, N., Hartline, J., and Malekian, A. 2012. Bayesian Optimal Auctions via Multi- to Single-agent Reduction. In *the 13th ACM Conference on Electronic Commerce (EC)*.

Alaei, S., Fu, H., Haghpanah, N., and Hartline, J. D. 2013. The Simple Economics of Approximately Optimal Auctions. In *the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*.

Anderson, E. J. and Nash, P. 1987. *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications*. John Wiley & Sons.

Babaioff, M., Immorlica, N., Lucier, B., and Weinberg, S. M. 2014. A Simple and Approximately Optimal Mechanism for an Additive Buyer. In *the 55th Annual Symposium on Foundations of Computer Science (FOCS)*.

Bhalgat, A., Gollapudi, S., and Munagala, K. 2013. Optimal Auctions via the Multiplicative Weight Method. In *the 14th ACM Conference on Electronic Commerce (EC)*.

Border, K. C. 1991. Implementation of Reduced Form Auctions: A Geometric Approach. *Econometrica 59,* 4, 1175–1187.

Border, K. C. 2007. Reduced Form Auctions Revisited. *Economic Theory 31*, 167–181.

Briest, P., Chawla, S., Kleinberg, R., and Weinberg, S. M. 2010. Pricing Randomized Allocations. In *the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*.

Cai, Y. and Daskalakis, C. 2011. Extreme-Value Theorems for Optimal Multidimensional Pricing. In *the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*.

Cai, Y., Daskalakis, C., and Weinberg, S. M. 2012a. An Algorithmic Characterization of Multi-Dimensional Mechanisms. In *the 44th Annual ACM Symposium on Theory of Computing (STOC)*.

Cai, Y., Daskalakis, C., and Weinberg, S. M. 2012b. Optimal Multi-Dimensional Mechanism Design: Reducing Revenue to Welfare Maximization. In *the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS)*.

Cai, Y., Daskalakis, C., and Weinberg, S. M. 2013a. Reducing Revenue to Welfare Maximization : Approximation Algorithms and other Generalizations. In *the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*.

Cai, Y., Daskalakis, C., and Weinberg, S. M. 2013b. Understanding Incentives: Mechanism Design Becomes Algorithm Design. In *the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*.

Cai, Y., Daskalakis, C., and Weinberg, S. M. 2015. Reducing Bayesian Mechanism Design to Algorithm Design. *Encyclopedia of Algorithms*.

Cai, Y. and Huang, Z. 2013. Simple and Nearly Optimal Multi-Item Auctions. In *the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*.

Chawla, S., Hartline, J. D., and Kleinberg, R. D. 2007. Algorithmic Pricing via Virtual Valuations. In *the 8th ACM Conference on Electronic Commerce (EC)*.

Chawla, S., Hartline, J. D., Malec, D. L., and Sivan, B. 2010. Multi-Parameter Mechanism Design and Sequential Posted Pricing. In *the 42nd ACM Symposium on Theory of Computing (STOC)*.

Chawla, S. and Sivan, B. 2014. Bayesian Algorithmic Mechanism Design. *ACM SIGecom Exchanges 13,* 1, 5–49.

Che, Y.-K., Kim, J., and Mierendorff, K. 2011. Generalized Reduced-Form Auctions: A Network-Flow Approach. *Econometrica 81,* 6, 2487–2520.

Chen, X., Diakonikolas, I., Paparas, D., Sun, X., and Yannakakis, M. 2014. The Complexity of Optimal Multidimensional Pricing. In *the 25th Symposium on Discrete Algorithms (SODA)*.

Daskalakis, C. Spring 2015. Lecture 21. In *Decision, Games, and Computation*. EECS, MIT. `https://stellar.mit.edu/S/course/6/sp15/6.891/index.html`.

Daskalakis, C., Deckelbaum, A., and Tzamos, C. 2012. Optimal Pricing is Hard. In *the 8th Workshop on Internet & Network Economics (WINE)*.

DASKALAKIS, C., DECKELBAUM, A., AND TZAMOS, C. 2013. Mechanism Design via Optimal Transport. In *the 14th ACM Conference on Electronic Commerce (EC)*.

DASKALAKIS, C., DECKELBAUM, A., AND TZAMOS, C. 2014. The Complexity of Optimal Mechanism Design. In *the 25th ACM-SIAM Symposium on Discrete Algorithms (SODA)*.

DASKALAKIS, C., DECKELBAUM, A., AND TZAMOS, C. 2015. Strong Duality for a Multiple-Good Monopoly. In *the 16th ACM Conference on Electronic Commerce (EC)*.

DASKALAKIS, C., DEVANUR, N., AND WEINBERG, S. M. 2015. Revenue Maximization and Ex-Post Budget Constraints. In *the 16th ACM Conference on Electronic Commerce (EC)*.

DASKALAKIS, C. AND WEINBERG, S. M. 2015. Bayesian Truthful Mechanisms for Job Scheduling from Bi-criterion Approximation Algorithms. In *the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*.

GIANNAKOPOULOS, Y. AND KOUTSOUPIAS, E. 2014. Duality and Optimality of Auctions for Uniform Distributions. In *the 15th ACM Conference on Electronic Commerce (EC)*.

HART, S. AND NISAN, N. 2012. Approximate Revenue Maximization with Multiple Items. In *the 13th ACM Conference on Electronic Commerce (EC)*.

HART, S. AND NISAN, N. 2013. The Menu-Size Complexity of Auctions. In *the 14th ACM Conference on Electronic Commerce (EC)*.

HART, S. AND RENY, P. 2012. *Maximal Revenue with Multiple Goods: Nonmonotonicity and Other Observations*. Center for the Study of Rationality.

HART, S. AND RENY, P. J. 2014. Implementation of Reduced Form Mechanisms: A Simple Approach and a New Characterization. *Economic Theory Bulletin 3,* 1, 1–8.

HARTLINE, J. D. 2013. Bayesian Mechanism Design. *Foundations and Trends in Theoretical Computer Science 8,* 3.

LI, X. AND YAO, A. C.-C. 2013. On Revenue Maximization for Selling Multiple Independently Distributed Items. *Proceedings of the National Academy of Sciences 110,* 28, 11232–11237.

LUENBERGER, D. G. 1968. *Optimization by Vector Space Methods*. John Wiley & Sons.

MANELLI, A. AND VINCENT, D. 2006. Bundling as an Optimal Selling Mechanism for a Multiple-Good Monopolist. *Journal of Economic Theory 127,* 1, 1–35.

MANELLI, A. M. AND VINCENT, D. R. 2007. Multidimensional Mechanism Design: Revenue Maximization and the Multiple-Good Monopoly. *Journal of Economic Theory 137,* 1, 153–185.

MYERSON, R. B. 1981. Optimal Auction Design. *Mathematics of Operations Research 6,* 1, 58–73.

PAVLOV, G. 2011. Optimal Mechanism for Selling Two Goods. *The B.E. Journal of Theoretical Economics 11,* 1 (February), 1–35.

RILEY, J. AND ZECKHAUSER, R. 1983. Optimal Selling Strategies: When to Haggle, When to Hold Firm. *The Quarterly Journal of Economics*, 267–289.

ROCHET, J.-C. 1987. A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context. *Journal of Mathematical Economics 16,* 2 (April), 191–200.

ROCHET, J.-C. AND STOLE, L. A. 2003. The Economics of Multidimensional Screening. *Econometric Society Monographs 35*, 150–197.

ROUGHGARDEN, T. 2015. Approximately Optimal Mechanism Design: Motivation, Examples, and Lessons Learned. *ACM SIGecom Exchanges 13,* 2, 4–20.

RUBINSTEIN, A. AND WEINBERG, S. M. 2015. Simple Mechanisms for a Subadditive Buyer and Applications to Revenue Monotonicity. In *the 16th ACM Conference on Electronic Commerce (EC)*.

TZAMOS, C. 2015. Optimal Mechanisms for Pairs of Beta Distributions.
`http://christos.me/betas.html`.

YAO, A. C. 2015. An n-to-1 Bidder Reduction for Multi-item Auctions and its Applications. In *the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*.