Mechanism Design for a Complex World: Rethinking Standard Assumptions

Kira Goldner

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:
Anna R. Karlin, Chair
Nikhil R. Devanur
Thomas Rothvoss

Program Authorized to Offer Degree:
Paul G. Allen School of Computer Science and Engineering

University of Washington

Abstract

Mechanism Design for a Complex World:
Rethinking Standard Assumptions

Kira Goldner

Chair of the Supervisory Committee:
Professor Anna R. Karlin
Computer Science and Engineering

The data used as input for many algorithms today comes from real human beings who have a stake in the outcome. In order to design algorithms that are robust to potential strategic manipulation, the field of algorithmic mechanism design formally models the strategic interests of the individuals and engineers their actions using game theory. The primary research directions in this area concern designing mechanisms to maximize either revenue or social welfare when selling to agents of various valuation types.

This thesis addresses barriers to progress in three fundamental directions in auction theory by rethinking standard models and assumptions and provides positive results in all three cases. First, we design revenue-optimal mechanisms in "interdimensional" settings—highly structured correlated settings that sit in between the assumed dichotomy of single-dimensional and multi-dimensional settings. Second, we propose a new model of proportional complementarities and construct an intuitive, simple mechanism that guarantees near-optimal revenue. Third, we study welfare maximization in the interdependent values setting without the single-crossing condition, and guarantee strong approximations for the most general setting of combinatorial auctions.

To Mom and Dad, for instilling independence and for giving me the freedom and encouragement to find my own passion and path.

And to Justin and Keith, my brilliant big brothers, for demonstrating the pursuit of passion and the self-discipline to succeed at it.

Acknowledgments

First and foremost, I must thank my advisor, Anna Karlin, for guiding me along the PhD journey. Your enthusiasm is contagious: for exciting research problems, for embarking on a new agenda, for making an impact, for working with new collaborators—and I have learned to appreciate and love all of these parts of research. Thank you for acting as the greatest role model, mentor, and teammate over the past five years. For teaching me lessons that I'm thankful every day to have learned: that it's better to admit confusion and use it as a learning opportunity, to prioritize intuition in presentations, but writing down math formally in research meetings, and to always be generous: with time, with credit, and even with sharing food. The principles you have instilled in me not only have improved my experience in academia, but make all of academia a better place.

I have been fortunate to have had a number of mentors throughout my PhD, people who have acted as role models, have hosted me at their institutions, and have developed and supported me as I have grown into the role of a researcher; I owe them great thanks. To Shuchi Chawla, for demonstrating how to fine-tune and formalize my half-baked ideas and writing during the earlier stages of my PhD. To Michal Feldman and Amos Fiat, for giving me a second home at Tel Aviv University, for their immense generosity with their time, and for conveying a new meaning of the word collaboration. To Nikhil Devanur, for his collaboration and guidance during and in addition to my internship, and for teaching me to always come at a problem with a fresh perspective (and perhaps some heavy technical machinery). To Brendan Lucier and Nicole Immorlica, for an internship that really stretched me in new directions, and for teaching me that the answer is always a geometric argument. And to Yang Cai and Matt Weinberg, whose work on mechanism design and appreciation of duality drew me to this research area in this first place, for your advice, support, and for letting me join your club. I am also very grateful to my dissertation committee: to Thomas Rothvoss and Nikhil for serving on my reading committee, as well as to Sebastien Bubeck and Jacques Lawarree for serving on my general dissertation committee.

I am thankful to my dearest friends, who support me every single day despite living far away. Thank you to Alon Eden for being the best collaborator and friend that I could ever dream of over the past few years. To Tselil Schramm, for guiding me toward great decisions and sharing genuine friendship and fits of laughter. To Matt Weinberg, for serving as an incredibly motivating benchmark, as well as an incredible mentor and friend. To Ben Miller, my first and long-term collaborator and friend, for your amazing patience: in listening, knowing me so well, and providing LaTeX assistance. To Rachel Cummings and Jamie Morgenstern, for ensuring that I treat myself and live my best life. To Yang Cai, for serving as a reference oracle, and for always letting me rant. And to Helen Jenne and Sam Petti, my randomly-assigned study-abroad roommates, who have been the most incredibly supportive friends ever since; I am so fortunate to have you both.

Thanks also to all of my collaborators throughout my PhD, past and ongoing: Anna, Alon, Amos, Ben, Brendan, Matt, Michal, Nicole, Nikhil, Shuchi, Yang, Moshe Babaioff, Meryem Essaidi, Yannai Gonczarowski, Elias Koutsoupias, Preston McAfee, Manolis Pountourakis, Raghuvansh Saxena, Ariel Schvartzman. Bouncing ideas back and forth with you has been the most enjoyable part of research. In particular, I owe thanks to my coauthors whose joint work with myself makes up this thesis. Chapter 3 is based on joint work with Amos, Anna, and Elias [FGKK '16]; Chapter 4 is based on joint work with Nikhil, Raghuvansh, Ariel, and Matt [DGSSW '19]; Chapter 5 is based on joint work with Yang, Nikhil, and Preston [CDGM '19]; Chapter 6 is based on joint work with Alon, Michal, Amos, and Anna [EFFGK '19].

Thank you to my friends at the University of Washington, particularly to Daniel Gordon and Eric Whitmire for their friendship over the past five years. To Amrita, Emily, Lucy, and Maarten, for their chats in the espresso room which maintained my sanity. Thanks to the whole UW theory group, especially to Becca Hoberg, Siva Ramamoorthy, Alireza Rezaei, and Robbie Weber for their solidarity.

There are too many other friends I should be thanking who have supported me, made me laugh, gossiped with me, and made this journey a joy—friends from academia, from undergrad, from childhood, and more. I do wish to call attention to the Tel Aviv University theory group, who made me feel so extraordinarily at home during my many visits, as well as Talya and Alon Eden (and their pupikim), who welcomed me into their family during my stays in Israel. I am also grateful to those who have embarked on the Mechanism Design for Social Good journey with me, especially Rediet Abebe for co-founding this initiative from the beginning, as well as Irene Lo for piling onto this crazy ride.

I also owe thanks to my mentors prior to my PhD who sent me toward this path.

To Alexa Sharp and Tom Wexler, I would not have found nor continued on the path of computer science without you. To Francine Blachet-Sadri, thank you for showing me the joys of research at a very impressionable time; I feel very lucky.

Finally, to my family, Mom, Dad, Justin, Keith, Zoe, and Cyrus, for their unending love and support. To Mommom, for always emphasizing education. And to Grandpa, for filling our lives with cards and music from childhood, and most especially, for all of your love.

Contents

1 Introduction

1.1 Designing Algorithms for Strategic Input

In the modern world, algorithms are increasingly used to make decisions that have tangible impacts on peoples' lives. Consider health insurance policies that determine which workers can seek treatment at which hospitals, or online platforms that determine which workers can apply for which jobs, or even a simple auction that determines which ad slots are allocated to which advertisers. When algorithms are designed under the assumption that their inputs will be accurate, but are instead run on (potentially manipulated) data that is produced by strategic individuals who have a stake in the outcome, the algorithm may perform poorly, and its guarantees may no longer hold.

Take, for example, Uber's algorithm for surge pricing. Their goal is to equalize driver supply and rider demand in order to maximize the number of successful driver-rider pairings. However, if drivers know that prices are likely to increase in the near future, or if their assigned ride turns out to be particularly low-paying, they may shut off their app in order to remain available for the higher-paying rides [Kelly, 2018]. When drivers do this, it misrepresents the true supply of drivers, and so the pricing and matching algorithms don't work as expected.

What is needed is a theory that accounts for the fact that if strategic agents can manipulate the input to the algorithm in order to improve their outcomes, then they will do so—a theory that can provide guarantees even given this sort of strategic input. This is where game theory comes in, providing tools to model and reason about the utility of strategic agents for every possible action that they might take.

The field of algorithmic mechanism design combines ideas from both algorithm design and game theory to produce *mechanisms*: algorithms that guarantee that, even when the participants act in their own self-interest, the designer's objective is achieved. *Game theory* is the study of strategic behavior where actions and payoffs are pre-defined; the challenge being to reason about which outcomes the strategic players will select. In contrast, *mechanism design* is sometimes referred to as "reverse game theory." Given knowledge about players' utility functions, the challenge is for the designer to choose the rules of the "game" in order to incentivize the players to behave in such a way that the designer's objective is met.

Algorithmic mechanism design merges ideas from computer science, such as notions of *tractability* and *approximation*, with economic notions, such as *incentives*, strategic behavior, and economic *efficiency*. Ideas at this intersection allow us to develop a theory that more closely aligns with the scenarios that modern day algorithms face when they interact with strategic users.

1.1.1 Overview of thesis

This thesis makes progress on three fundamental problems in auction theory by rethinking standard models and assumptions.

Interdimensional mechanism design. The problem of designing revenue-maximizing auctions, also known as *optimal* auctions, has been the subject of intense study for decades. The problem is completely solved when the auctioneer has a single item to sell [Myerson, 1981]; however, even in the case where a seller has only two items to sell, we fail to fully understand optimal auctions. The optimal auction in this setting can be incredibly complex—highly randomized and intractable to compute—and thus characterizations remain elusive.

Our work introduces a new subclass of natural and ubiquitous multi-dimensional settings that we call *interdimensional* settings. The complexity of maximizing revenue in these settings lies strictly between the "easy" case of "single-parameter" settings (where the buyer's private information is characterized by a single number) and the fully general "multi-dimensional" setting. For an example, consider the case where a seller offers items that vary in levels of quality.

Our main result is a complete characterization of the optimal solution in the interdimensional setting we call "The Fedex Problem." This is one of the first multi-dimensional settings where a closed-form solution is obtained without any restriction on the form of the mechanism and without any assumption about the prior distributions from which buyer values are drawn.

We then characterize the optimal mechanism for selling to a single-minded buyer. While this is only a slight generalization of the FedEx setting, we find that the *menu complexity*, or the number of distinct options offered to the buyer, jumps from exponential in the FedEx setting (specifically, 2^{m-1} for m services) to *unbounded but finite* (even for $m = 3$ services) in the single-minded setting. This sharply separates the single-minded setting from both the FedEx setting and the "multi-dimensional" two-item additive setting, which is known to have uncountably infinite menu complexity [Daskalakis, Deckelbaum, and Tzamos, 2017].

Mechanism design with complements. Items have *complements* when a buyer derives some additional value from receiving a set of items beyond just the sum of his values for the individual items. A classic example concerns a left and right shoe: obviously, a buyer derives significantly more value for a pair of shoes than the sum of his value for a lone left shoe plus his value for a lone right shoe.

Positive results for maximizing revenue in settings with complements have been few and far between, and either preclude decent approximation factors or rely on the assumption of independence between valuations for every different bundles of items (which clearly does not always make sense, e.g., consider the left shoe, the right shoe, and the pair of shoes).

Our work revisits the question of how to model complementarities. We introduce a new model that replaces the independence assumptions of previous models with known proportionalities observed from data. Our model is motivated by the Microsoft Pricing Engine, a product that was developed to take in a company's sales history dataset and output a suggestion for what prices to sell that company's products at, whether to sell them individually or bundle them together, and how to capitalize on any complementarities that may exist among the items. For example, if it is known that users derive some extra value from using Microsoft Word with Excel, say, by creating charts in Excel and dragging them into Word, it may increase the seller's revenue to sell Word and Excel together as a bundle.

Our main result is a new, simple, and intuitive mechanism that capitalizes on these proportionalities. As a result, we are able to give approximations to revenue that are linear in the smaller of (1) the maximum-degree and (2) the largest-hyperedge for a hypergraphic model of complementarities. This contrasts sharply with prior work which gives an approximation that is linear in the maximum-degree but exponential in the largest-hyperedge, and also requires a strong independence assumption.

Mechanism design with interdependent valuations. In the standard model of *independent private values*, a buyer's value for an item is unaffected by other buyers' values for the item. This is often unrealistic. For an item such as a house or a painting, the item's resale value is directly relevant to the buyer's willingness-to-pay, and is well-approximated by the values of other buyers. Further, in many cases, buyers do not actually know their values for an item, but rather, each have partial information regarding the item. Combining the information of all buyers informs a buyer of his value for the item. For example, if two firms are bidding for oil drilling rights, and one firm learns the amount of oil available, that information directly impacts the value that the other firm has for the drilling rights. The *interdependent values model* [Milgrom and Weber, 1982] captures the idea that a buyer's information and value may impact the values of other buyers.

Unfortunately, for interdependent valuations, there are strong impossibility results that preclude the existence of mechanisms for attaining optimal social welfare beyond very restricted settings (such as single-item auctions when buyer valuations satisfy the "single-crossing" condition).

Our work introduces a new assumption on interdependent valuations that we call "submodularity over signals." This assumption is natural and benign; in particular, it is satisfied by the typical examples of interdependent valuations in the literature. With this assumption in hand, we obtain the first constant-factor approximations to social welfare for combinatorial auctions, the most general auction environment, in the interdependent values model.

In the rest of this chapter, we present a brief overview of results in each of these three directions, with formal definitions deferred to chapters 3-6.

1.2 Interdimensional Mechanism Design

In this section, we sketch the landscape of known revenue-optimal mechanisms, highlighting the sharp contrast between single-dimensional and multi-dimensional settings. We then overview our results for the FedEx setting and the single-minded setting, and discuss the features of these interdimensional settings that provide traction for a duality approach. Full details are provided in Chapters 3 and 4.

1.2.1 Background: Optimal Revenue

One of the most fundamental problems in mechanism design concerns an auctioneer who owns m items and wishes to design a revenue maximizing auction for selling them. In the standard model for this setting, there are n potential buyers, and each buyer i has a value for each item j. From the seller's perspective, buyer i's value for item j is drawn from an independent prior distribution F_{ij}, of which the seller has full knowledge. What mechanism for selling the m items should the seller implement in order to extract the most revenue possible (in expectation over the prior distributions)?

As we will discuss in Chapter 2, by the revelation principle, it is without loss to restrict our attention to *truthful*—or (Bayesian) *incentive compatible*—mechanisms. That is, we can focus on mechanisms for which it is in the best interest of every buyer to report his true valuations for each item to the mechanism.

If $m = 1$, then the seller only has a single item to sell, and this problem is completely resolved with a beautiful closed-form theory [Myerson, 1981; Riley and Zeckhauser, 1983; Bulow and Roberts, 1989]. Each buyer has a *virtual value function* $\varphi_i(\cdot)$ mapping values to so-called "virtual values". Myerson shows that the expected seller revenue is maximized by allocating to the buyer i with the highest virtual value $\varphi_i(v_i)$, that is, maximizing expected *virtual welfare*. As this process only yields a truthful mechanism for *regular* buyer distributions, Myerson defines an *ironing* procedure to accommodate *irregular* distributions. We define all these concepts formally in Section 2.3.

Myerson's virtual value theory extends to all *single-parameter* environments. This includes, for example, selling k identical items to buyers who wish to buy one item. Extending Myerson's theory or introducing an analogue for *multi-dimensional* environments, where a buyer has more than a scalar for private information, has been a long-standing open problem in mechanism design.

Unfortunately, optimal multi-dimensional mechanism design proves to be exceptionally challenging. For example, the optimal mechanism to sell two items to a single additive bidder might offer uncountably infinite options to the buyer [Daskalakis et al., 2017; Manelli and Vincent, 2006], while no mechanism that offers only a bounded number of options can guarantee a finite approximation to the optimal revenue [Briest, Chawla, Kleinberg, and Weinberg, 2015; Hart and Nisan, 2017]. As a result, most work aimed at this goal has instead either characterized the optimal mechanism for restricted settings or distributions [Laffont, Maskin, and Rochet, 1987; McAfee and McMillan, 1988; Giannakopoulos and Koutsoupias, 2014, 2015], or has provided sufficient conditions under which a particular mechanism is optimal [Haghpanah and Hartline, 2015; Daskalakis et al., 2017]. See Section 3.1.1 for an in-depth overview.

1.2.2 Our Contribution: Optimal Results in Interdimensional Settings

As we have seen, there is quite a dichotomy painted between "single-dimensional" settings (e.g. where all items are identical) and "multi-dimensional" settings (e.g. where all items are heterogenous). Single-dimensional settings have optimal auctions which are described by a beautiful closed-form theory, offer only one non-trivial option, and are simple to compute. In contrast, multi-dimensional settings have optimal auctions which may offer infinitely many options, can be intractable to compute [Daskalakis, Deckelbaum, and Tzamos, 2013], and in addition to many other undesirable properties, elude characterization thus far.

In Chapters 3 and 4, we study optimal auction design in a regime that is fundamentally in between. Specifically, we rethink the standard assumption that settings with multiple items must either be identical or completely heterogenous with independently drawn valuations. Instead, we introduce settings such as in the following scenario, where varying qualities of service are offered and the customer's value is highly correlated across these services.

Scenario 1: The FedEx Problem. A customer has a package to ship and a deadline that he needs the package received by. If he purchases a shipping option that will deliver the package on or before his deadline, then the shipping service is worth some value v to him, otherwise it is worth nothing. The (value, deadline) pairs are drawn from an arbitrarily correlated joint distribution F. The seller, FedEx, sells a variety of shipping options: 1-day, 2-day, and so on, up to m-day shipping. Given the prior F, what mechanism maximizes FedEx's expected revenue?

Observe that the buyers are unit-demand for one of m different shipping options with a specific correlated structure over their values for the options. In particular, if a customer has a value v for receiving a package by deadline d, then he values 1-day, 2-day, and up to d-day shipping at v each, and $(d+1)$-day through m-day shipping at 0. This leads us to our first research question, which we address in Chapter 3.

Question 1. What is the revenue-optimal mechanism for selling to a single "FedEx" buyer, whose prior distribution is known?

What is special about the FedEx setting is that if we consider only the customers with a deadline i days from now, it's as if we have a single-dimensional problem. These single-dimensional problems, one for each potential deadline, are stitched together by the inter-day incentive compatibility constraints, creating a very specific type of multi-dimensional problem. As we will see, the incentive compatibility constraints within each deadline, as well the constraint that each type prefers his deadline to reporting one deadline earlier, are enough to ensure global incentive compatibility (IC). This limited number of relevant constraints also means that we can use Myerson's payment identity. We depict these IC constraints in Figure 1.1.

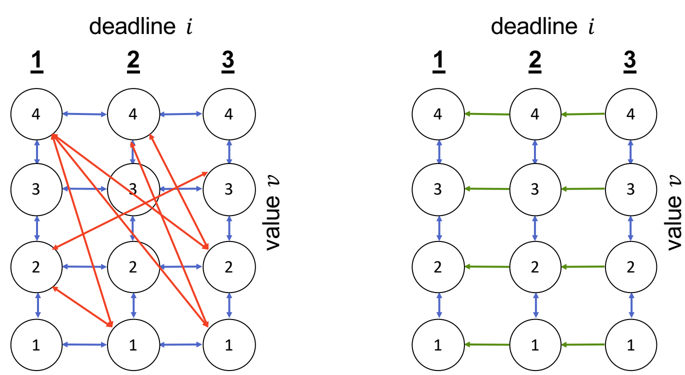


Figure 1.1: A depiction of the incentive compatibility constraints in FedEx for a discretization of the value space (for ease of visualization). An arrow from (v,i) to (v',i') indicates that type (v,i) must prefer truthful reporting to misreporting type (v',i'). The left illustration contains the analogue of single-dimensional constraints that we would hope for in the blue grid, as well as the red constraints that we expect to have for all pairs of types. However, the structure of the FedEx setting enables a large reduction in IC constraints, as depicted on the right. We only need that a type prefer reporting their value over any other value for a fixed deadline (the vertical blue constraints, replaced by the payment identity) and for a type to prefer reporting their deadline over one deadline earlier (the asymmetric green constraints).

We attack the problem using a duality approach similar to that of Giannakopoulos and Koutsoupias [2014]. We leverage these two features of the FedEx setting—the limited number of IC constraints, and the use of the payment identity to swap out payment variables for constraints on allocation variables—to decrease the number of variables and constraints in the primal and dual problems such that we are able to come up with a solution. These optimal variables provide a characterization of the optimal mechanism for *any* prior distribution over (value, deadline) pairs. This is one of the first results of this form beyond single-parameter settings, and is in contrast to the prior optimal revenue results in multi-dimensional settings, which are either for restricted distributions or particular mechanisms.

The optimal mechanism takes the following form. Each shipping option will have a price—possibly randomly drawn from a distribution of prices, of which the buyer has full knowledge for each shipping option. The buyer will choose the shipping option whose price, in expectation over the randomization, maximizes his utility, and will commit to using this shipping option alone. Because the mechanism is incentive compatible, if the buyer's deadline is in d days, he will commit to d-day shipping. Then, the price for d-day shipping, p_d, will be drawn randomly from the distribution. If the buyer's value $v < p_d$, then he will not buy shipping and will not ship his package, but he cannot now purchase a different shipping option. If $v \geq p_d$, then he will ship his package with d-day shipping. The prices are determined by a combined-deadline revenue curve, which acts like a dynamic program for the revenue for all later days. In fact, these combined revenue curves are precisely the optimal dual variables.

Further, our characterization parallels Myerson's single-item characterization in some ways, with appropriate modifications to handle the more complex setting. For example, we optimize these combined revenue curves, which require a form of Myerson-like "ironing," except in value-space instead of quantile-space[1], in order to ensure that the inter-day incentive-compatibility constraints align. By studying this structured "interdimensional" realm, we are able to develop a closed-form solution, while side-stepping many of the negative results of "truly" multi-dimensional settings, such as the infinite menu complexity result for correlated distributions [Daskalakis et al., 2017].

Further, combining our approach with the Lagrangian approach of Cai, Devanur, and Weinberg [2016], enables us to view our solution through the language of *virtual*

[1]Quantile space corresponds to value space using $q = 1 - F(v)$ for any value v. Note that $q \in [0, 1]$.

welfare maximization, and to describe precisely what the analogue of Myersonian *virtual values* are in the FedEx setting.

Single-Minded Agents. The FedEx setting is not the only environment possessing the features that give duality much more traction toward characterizing the optimal mechanism. In fact, one way to view the "items" for sale (shipping options) in the FedEx setting is as a totally-ordered set where i-day shipping is better than $(i+1)$-day shipping. A generalization of this setting is selling partially-ordered items, or equivalently, selling bundles to single-minded bidders. For example, an internet service provider might be selling internet, a bundle with internet/phone, and a bundle with internet/cable. A customer may show up with an interest in any of these three bundles. Clearly, internet/phone and internet/cable are both at least as good as internet, however internet/phone and internet/cable are incomparable to each other. That is, these bundles are partially-ordered where the relation is set inclusion. This is depicted in Figure 1.2.

Figure 1.2: A depiction of the smallest example of a single-minded instance.

This introduces our second research question, which we address in Chapter 4.

Question 2. What is the revenue-optimal mechanism for selling to a single single-minded (or "partially-ordered") buyer, whose prior distribution is known? How does the optimal mechanism vary in complexity from FedEx, single-dimensional, and multi-dimensional settings?

Extending the approach for the FedEx setting to this problem and others, we provide a characterization of the optimal mechanism, giving lower and upper bounds on the degree of randomization needed, and developing a deeper understanding of the use of ironing. We combine the Lagrangian approach of [Cai et al., 2016] with the duality framework of [Giannakopoulos and Koutsoupias, 2014] to come up with a new duality approach for characterizing optimal mechanisms.

We show that these "interdimensional" settings fall on the spectrum between single- and multi-dimensional by several metrics: (1) how precisely we can describe the optimal mechanism, (2) the degree of randomization required, and (3) the conditions that imply that the optimal mechanism is deterministic.

The Class of Interdimensional Settings. Many natural settings, such as the FedEx setting or the Single-Minded setting, turn out to be interdimensional. They fill in the spectrum that was previously assumed to be a dichotomy between single- and multi-dimensional settings, and give some of the first unrestricted progress on characterizing optimal mechanisms in (limited) multi-dimensional settings. Since our work in the FedEx setting, numerous follow-up works have studied interdimensional settings, characterizing the revenue-optimal mechanisms or analyzing their menu complexity. Saxena, Schvartzman, and Weinberg [2018] study the menu complexity of $(1-\varepsilon)$-approximations to revenue in the FedEx setting. Devanur and Weinberg [2017] study the Budgets setting, where a buyer has a value for an item and a private budget b which is the most that he can pay. Devanur, Haghpanah, and Psomas [2017] study the Multi-Unit Pricing setting, where a buyer has a value for a unit of an item, and a private demand capacity d, and he is additive for each unit he receives up to his demand capacity. That is, his value for k units is $\min\{k, d\} \cdot v$. In each of these settings, the incentive compatibility constraints between types are more limited than "fully multi-dimensional" settings, but the seller still must carefully reason about them to design the optimal mechanism. See Section 3.7 for more details about this class of problems.

1.3 Simple and Approximately Optimal Pricings with Complements

In this section, we describe the limitations of prior work on settings with complements and our approach to circumventing these limitations. We begin with an overview of a common workaround to the complexity of optimal mechanism design: the use of mechanisms that are both simple and approximately optimal. We then describe a new model of proportional complementarities and the simple and approximately optimal pricing that we construct for this model. Details are provided in Chapter 5.

1.3.1 Background: Simple and Approximately Optimal Pricings

As we have discussed, in multi-dimensional settings, the precise optimal mechanism tends to be complex and elusive in many ways. This has motivated research on the design of mechanisms that are *simple*, but still provide near-optimal expected revenue guarantees.

One of the first such results was by Chawla, Hartline, Malec, and Sivan [2010], who consider the setting of n *unit-demand* buyers, who each want at most one of m possible items, and have values that are drawn from independent priors. They prove that selling the items separately gives a 30-approximation to the optimal revenue. Another result of this type is for a single *additive* buyer with values drawn independently across m items. Babaioff, Immorlica, Lucier, and Weinberg [2014] show that the better of (1) selling separately (by posting the reserve price on each item) and (2) selling the grand bundle (by posting the reserve price for the distribution of the sum of all item values) gives a 6-approximation to the optimal revenue in expectation. This is surprising, since neither selling separately nor grand bundling on their own can guarantee a constant-factor approximation [Hart and Nisan, 2017]. This result was extended, with modifications in the mechanism, for a single subadditive buyer [Rubinstein and Weinberg, 2015], multiple additive buyers [Yao, 2015], multiple constrained-additive buyers [Chawla and Miller, 2016], and multiple XOS and subadditive buyers [Cai and Zhao, 2017]. These results were also unified in a framework by Cai et al. [2016]. For more details, see Section 5.1.1.

All of the above simple-and-approximately-optimal results apply only in settings where, (1) there is independence between items and (2) at minimum, subadditivity holds. This assumes that a buyer's valuations are complement-free. However, there are many settings where a buyer derives some extra value from owning a combination of items that is not present from owning any item individually; that is, the items have *complementarities*.

The first work focusing on revenue maximization for complements is that of Eden, Feldman, Friedler, Talgam-Cohen, and Weinberg [2017b], who aim to find a simple and approximately-optimal mechanism for a single buyer with complementarities given by the MPH-k model, introduced as the PH-k model by Abraham, Babaioff, Dughmi, and Roughgarden [2012] and extended to the MPH-k model by Feige, Feldman, Immorlica, Izsak, Lucier, and Syrgkanis [2015]. In this model, a buyer's valuation is given by (the maximum over) a weighted hypergraph, and a buyer's valuation for any set of items S is then the sum over his value for the activities associated with all subsets of items $T \subseteq S$ (the weighted edges w_T), that is,

$$v(S) = \max_{\ell} v^{\ell}(S) \qquad \text{where} \qquad v^{\ell}(S) = \sum_{i \in S} w_i + \sum_{T \subseteq S} w_T.$$

To utilize the framework of previous results, the authors assume that the buyer's hyperedge types are drawn independently from known distributions. Their result is that the better of selling separately and the grand bundle in the MPH-k model gives an $O(d)$-approximation, where d is the largest degree of any vertex, or the largest number of edges that any one item appears in. Further, they prove that this factor d is necessary. In addition, they show that the approximation is exponential in the positive rank k, the size of the largest hyperedge, despite the fact that the welfare approximations from Abraham et al. and Feige et al. are linear in k.

1.3.2 Our Contribution: Proportional Complements

We begin with a motivating example for our new model:

Scenario 2: Selling Microsoft Office. Microsoft is selling Office products, including Word and Excel, and these items have *complementarities* among them. A buyer has some value for owning Word and some value for owning Excel. However, if he owns Word and Excel together, he derives some extra value from the pair that was not present from owning either product individually; this corresponds to the benefit of owning the two products together, such as creating a chart in Excel and dragging it into a Word document. The amount that this benefit increases the buyer's willingness-to-pay for the pair of both Word and Excel is proportional to the buyer's values for Word and Excel individually. Given the market parameters that determine these proportionalities, how should Microsoft sell Office products in order to (approximately) maximize revenue?

In the model of Eden et al., the buyer's value for {Word, Excel} would consist of his values for activities of Word, Excel, and Word + Excel, and his draws for these three values would be *independent*. Should he have higher values for Word and

Excel individually, this would have no bearing on his value for the activities he can do with Word + Excel together. This seems very unrealistic. However, without the independence assumption from Eden et al. [2017b], strong negative results for correlated distributions hold [Briest et al., 2015; Hart and Nisan, 2017], and no finite approximation to revenue is possible.

In Chapter 5, we propose a new model where a buyer has base valuations for each item, and the buyer's value for a set of item's are *proportional* to his base valuations for the individual items. In this model, the complementarities among items are parameters of the market. This is particularly relevant in a case such as selling Microsoft Office products, where the items are used together in an approximately fixed way, and the seller is able to collect data on how they are used and estimate, e.g. via cross-price elasticities, these market parameters.

In our model, for each item i, the buyer has a base valuation t_i, and these base valuations are drawn from independent distributions. Then, for an item i and a disjoint set of items T, a market parameter η_{iT} describes the boost from receiving the bundle that includes all the items in set T in addition to item i: the buyer derives an additional value of $\eta_{iT} t_i$ from item i. Thus, if a buyer receives a set S, his value for item i is

$$\eta_i(S)t_i := \left(1 + \sum_{T \subseteq S \setminus \{i\}} \eta_{iT}\right) t_i$$

where the 1 represents his value for item i itself, and his value for the set of items S is

$$v_i(S) := \sum_{i \in S} \eta_i(S) t_i.$$

Notice that the *boosts* that the buyer receives from owning additional items are described asymmetrically here, which is only more general. It is easy to instantiate symmetric boosts.

Example 1. Suppose that the three items are Powerpoint, Excel, and Word. The market parameters η that determine a buyer's boosts are given in the form of a hypergraph, as depicted in Figure 1.3. If the buyer were to receive Powerpoint and Excel, then his valuation for this bundle would be: his base valuation for Powerpoint (t_{ppt}) plus the boost onto Powerpoint he gets from also owning Excel $\eta_{\text{excel}\to\text{ppt}} \cdot t_{\text{ppt}}$, plus his base valuation for Excel (t_{excel}) plus the boost onto Excel he gets from also

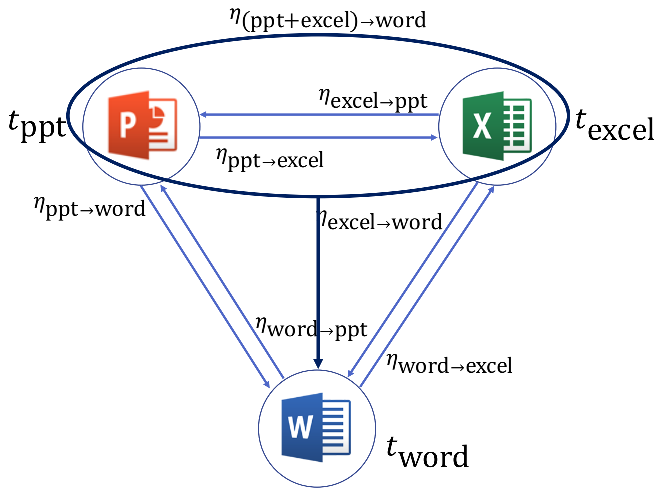


Figure 1.3: A depiction of the proportion complements model instantiated for three Microsoft Office products.

owning Powerpoint $\eta_{\text{ppt}\to\text{excel}} \cdot t_{\text{excel}}$. Note that we describe this second quantity as $\eta_{\text{excel}}(\{\text{excel} + \text{ppt}\}) \cdot t_{\text{excel}}$. Should the buyer receive all three items, then he will also get a boost from owning both Powerpoint and Excel onto Word, $\eta_{(\text{ppt+excel})\to\text{word}}$.

Our primary question is as follows.

Question 3. Can a simple mechanism guarantee a constant-factor approximation to the optimal revenue? Or, at the very least, an approximation linear in the parameters d or k of the hypergraph? How can we use the specific knowledge of the market parameters η in the mechanism?

Unsurprisingly, we find that selling items separately at their individual reserve prices is not a good choice in this setting because of the complementarities between the items. Instead, we find that the better of two mechanisms, selling the grand bundle and running a new mechanism we call SEPARATE/FREE, gives a $O(\min\{d, k\})$-approximation to revenue. Contrast this with Eden et al. [2017b], whose approximation is linear in d but exponential in k[2].

[2]In the complete hypergraph with hyperedges of size $k = \log n$ (and the maximum degree $d = \binom{n}{\log n}$), our mechanism guarantees an $O(\log n)$-approximation, while Eden et al. [2017b] guarantees at best an $O(n)$-approximation. Of course, the models are quite different.

The mechanism SEPARATE/FREE first gives a number of items away for free, which the buyer will, of course, happily take. Then, capitalizing on the fact that there are complementarities among these free items and the items remaining for sale, the latter (which are now more valuable to the buyer) can be sold separately at inflated prices. We see this mechanism often in practice, such as Android giving the operating system away for free to sell ads across it, or Microsoft giving OneNote away for free with compatibility to Office.

We also use the [CDW '16] Lagrangian Duality framework in a new way. The standard approach would derive an upper bound via duality in the complements setting and cover it directly with mechanisms. Instead, we upper bound the revenue in the complements setting by the revenue in an inflated setting, and our proof is via the optimal dual variables, covering the variables in the complements setting with those from the inflated setting. Then, we use the standard approach on the inflated setting, and bound this revenue with mechanisms from the complements setting.

1.4 Welfare Maximization for Interdependent Valuations

In this section, we highlight the difficulty of maximizing social welfare in the interdependent value setting. We overview our mechanism design goals, introduce the natural *submodular over signals* assumption, and summarize our results. Full details appear in Chapter 6.

1.4.1 Background: Limited Results for Restricted Interdependent Valuations

A common goal in mechanism design is to maximize social welfare (or economic efficiency)—that is, to allocate the goods to the bidders who value them the most. In the typical mechanism design model, each buyer has a *private* value for each item being sold, and the fact that buyer 1 values a coffee at $3 has no impact on buyer 2's value for the same coffee. For private values, no matter what type of valuations buyers have for the goods (e.g. complements or substitutes), the Vickrey-Clarke-Groves (VCG) mechanism truthfully maximizes welfare [Vickrey, 1961; Clarke, 1971;

Groves, 1973]. The VCG mechanism finds the allocation of items to buyers that maximizes welfare and computes appropriate payments to make this allocation truthful.

In contrast to private values, we study the interdependent values model, depicted in the following scenario.

Scenario 3: Buying a House. A number of customers are interested in purchasing a house. Each potential buyer has an inspection performed on the house which returns some piece of private information: perhaps agent 1 learns about the quality of the foundation, while agent 2 learns about the quality of the plumbing, agent 3 the electrical, agent 4 the windows, and so on. Any agent's value for the house depends on *all* of these pieces of private information, even though they are unknown to the agents. Thus, agents do not know their own value for the house; rather they know their valuation as a *function* of the unknown private information, in this case, the quality of the plumbing and of the electrical. How can the seller solicit the private information in order to compute these valuation functions and determine how to allocate the house to (approximately) maximize social welfare, e.g. giving the house to the buyer who (approximately) values it most?

Formally, each buyer i has some private information s_i, which we call his *signal*, and a valuation function $v_i(s_1, s_2, \ldots, s_n)$ which depends not only on his own private information, but on the private information of possibly all n buyers. In this event, a buyer may not even know his own value, only that it increases at a certain rate with the private information of another buyer, whose private information he does not know. This complicates incentives a good deal. Notice that in the interdependent values model that we use, initiated by Milgrom and Weber [1982], the form of each buyer's valuation function is assumed to be public, although the inputs to this function are private, and distributed across the buyers. Take the following scenario from Dasgupta and Maskin [2000].

Example 2 ([Dasgupta and Maskin, 2000])**.** Consider two wildcatters who are competing for the right to drill for oil on a given tract of land. The wildcatters' costs of drilling differ. Wildcatter 1 has a fixed cost of 1 and a marginal cost of 2 (per unit of oil extracted). Wildcatter 2's fixed cost is 2 and marginal cost is 1. Oil can be sold at a price of 4. Wildcatter 1 performs a (private) test and discovers that the expected size of the oil reserve is s_1 units. Wildcatter 2's private information s_2 does not affect either driller's payoff. We have

$$v_1(s_1, s_2) = (4 - 2)s_1 - 1 = 2s_1 - 1$$

and

$$v_2(s_1, s_2) = (4 - 1)s_1 - 2 = 3s_1 - 2.$$

Now, even though the buyers do not necessarily know their valuations, we could try to use the *Generalized VCG* (G-VCG) mechanism, proposed in Dasgupta and Maskin [2000]. In this generalization of VCG, the bidders are first asked to report their signals (rather than values) to the auctioneer. Since the valuation functions are publicly known, the auctioneer can then compute all of the buyers' valuation functions on the reported signals. If the reported signals are truthful, then the auctioneer outputs the welfare-maximizing allocation, and tries to reverse-engineer payments that will incentivize truthfulness (modified for signal space instead of value space).

However, there is a problem with this proposal in Example 2. Suppose we use the G-VCG auction and just give the drilling rights to the firm with the higher value. If the first firm discovers that the reserve size is large, then the second firm will have a much larger value, and thus will win the drilling rights. Hence, the first firm has the incentive to misreport his signal and claim that the oil reserve is smaller than it actually is. This will mislead the auctioneer into believing that firm 1's value is larger than firm 2's. In this way, the interdependent value setting allows a buyer not only to misreport his own private information, but to corrupt beliefs about other buyers' valuations as well.

The above G-VCG mechanism is sometimes truthful—when the valuations satisfy what is called the *single-crossing condition*. This condition states, in essence, that a buyer is more sensitive to his own signal than any other buyer, and rules out precisely the type of scenario we see in Example 2. Under the single-crossing condition, the G-VCG mechanism attains optimal social welfare for some single-parameter settings under matroid feasibility constraints, and thus welfare maximization under single-crossing for many important single-parameter settings is solved. On the other hand, without single-crossing, there exists valuations such that the allocation that achieves optimal social welfare cannot be implemented with a truthful mechanism. For this reason, almost all prior work in interdependent value settings assumes the single-crossing assumption.

In addition, almost all work in the interdependent setting supposes that each bidder i has a single *signal*, or piece of private information, s_i. We refer to this as the case where signals are "single-dimensional."

The two most relevant works that consider interdependent settings where bidders instead have multiple or "multi-dimensional" signals are Dasgupta and Maskin [2000] and Jehiel and Moldovanu [2001]. These works largely consist of impossibility results, proving that when a bidder has multiple signals per outcome *or* one signal for each of multiple possible outcomes, unless the valuations are degenerate, the allocation that maximizes social welfare does not satisfy even Bayesian incentive compatibility. These works also contain limited positive results, but contain *very* strong assumptions, such as: a separability condition which essentially compresses the signals to a single-dimensional statistic or valuation functions that are linear functions of the signals.These two papers also focus on *Bayesian incentive compatibility*, which requires that buyers' signals are drawn from prior distributions that are known to *both* the auctioneer and other buyers. See Section 6.1.4 for further details.

1.4.2 Our Contribution: Beyond Single-Crossing

Our goal in Chapter 6 of this thesis is to maximize welfare *without* the single-crossing assumption. We know that attaining the optimal welfare without single-crossing is impossible; instead, we aim to *approximate* the optimal welfare. The first attempt at approximation without single-crossing was made in Eden, Feldman, Fiat, and Goldner [2018], where strong lower bounds for approximation are shown when the valuations do not satisfy single-crossing and are otherwise unrestricted. These lower bounds imply that to guarantee good approximations, we must restrict the valuation class in *some* way. Eden et al. [2018] investigate a parameterized form of single-crossing, *c-single-crossing*, and obtain a deterministic $O(nc)$-approximation and randomized $O(\sqrt{n}c^{3/2})$-approximation.

Another objective of our work is to eliminate the dependence of the results on priors. Thus, we ask that our equilibrium notion be stronger than Bayesian incentive compatibility.[3] Specifically, we seek ex-post incentive compatible mechanisms that

[3] The strongest notion we might try to achieve, dominant-strategy incentive compatibility, is not possible to enforce in interdependent settings due to the fact that buyers have the potential to corrupt other buyers' valuations by misreporting their signals

obtain strong prior-free welfare guarantees. *Ex-post incentive compatibility* requires that reporting truthfully be a best response to all other agents reporting their true signals, and *prior-free* guarantees require that the approximation holds for every realization of signals.

Our third goal is to extend beyond the realm of single-dimensional signals and single-item (or single-parameter) settings. Ideally, we would like our mechanisms to work well in settings as broad as combinatorial auctions, where a buyer may have a value for every possible bundle of items, and might also have a signal for every bundle.

All together, these goals culminate in the following research question.

Question 4. What truthful, prior-free approximation to welfare can we guarantee without single-crossing? For single-dimensional or multi-dimensional signals? For single-parameter settings? Beyond single-parameter settings? For combinatorial auctions?

And, since we must restrict the valuation class to avoid the lower bound, we naturally must ask:

Question 5. What natural, not-too-restrictive condition on the valuation class allows strong approximation guarantees?

In Chapter 6, we show that if the valuation functions are *submodular over signals*, then we can guarantee very strong approximations. We can interpret this condition as meaning that when less information is available, e.g., when some signals are fixed at lower values, buyers are more sensitive to a change in some other signal. This condition arises naturally in many examples from the literature.

Submodularity allows us to use a simple random-sampling idea in our mechanisms to help unentangle some of the incentive problems. We split the buyers into two sets at random, A and B. B contains our "potential winners," and we do not use the signals in B to determine the allocations to other potential winners; we only use the signals of non-winners from A. That is, we construct "proxy valuations": for $i \in B$, we consider i's valuation computed given the reported signals in A, i's own signal, and none of the other signals in B. For $i \in A$, i's proxy valuation is 0. We then select the allocation that maximizes welfare with respect to the proxy valuations. The key observation is that submodularity ensures that our proxy valuations are close enough in expectation to the true values such that they can be used as a good approximation for welfare maximization.

This technique, combined with a new way to set payments, allows us to move beyond the realm of single-parameter settings, and to obtain strong approximately-optimal welfare guarantees in more general interdependent settings such as unit-demand settings and combinatorial auctions. These are the first positive results in these environments for interdependent values, in part because single-crossing or generalizations thereof does not suffice in these environments for the maximum welfare allocation to be attainable by a truthful mechanism.

In addition to submodularity and the random-sampling mechanism expanding the terrain of positive results all the way to combinatorial auctions, our guarantees also improve the best known approximation without single-crossing from $O(\sqrt{n}c^{3/2})$ under c-single-crossing to 4 under submodularity, suggesting that submodularity may be the correct condition for approximation.

1.5 Bibliographic Notes

Chapter 3 is based on joint work with Amos Fiat, Anna Karlin, and Elias Koutsoupias [FGKK '16], which appeared at EC 2016. Chapter 4 is based on ongoing joint work with Nikhil Devanur, Raghuvansh Saxena, Ariel Schvartzman, and Matt Weinberg [DGSSW '19]. Chapter 5 is based on joint work with Yang Cai, Nikhil Devanur, and Preston McAfee [CDGM '19], which appeared at EC 2019. Chapter 6 is based on joint work with Alon Eden, Michal Feldman, Amos Fiat, and Anna Karlin [EFFGK '19], which appeared at EC 2019, and was awarded Best Paper with a Student Lead Author.

2 Preliminaries

In these preliminaries, we will define mechanism design concepts with respect to typical assumptions, and in the proceeding chapters, we will modify the definitions as needed as we move beyond various assumptions.

Section 2.1 defines the basics of a mechanism design setting. Section 2.2 discusses constraints pertaining to incentives and when and how to implement them. Section 2.3 overviews revenue maximization preliminaries. Section 2.4 provides an introduction to Lagrangian duality. The experienced reader may wish to skip this chapter or simply refer to it as an appendix.

2.1 Mechanism Design Basics

For the purpose of this thesis, an auction contains n buyers or bidders, with bidder $i \in \{1, \ldots, n\}$. An auctioneer sells m items, denoted $j \in \{1, \ldots, m\}$. Any player that is strategic is referred to as an *agent*, hence buyers are also agents.

Types. A buyer has a type v that is his private information. What this type represents precisely varies depending on the setting at hand.

In a single-item setting, there is an item for sale, and v is a scalar that represents the buyer's *value*, his maximum willingness-to-pay for the item, or how much the item is worth to him. If he pays $v - 3$ dollars to receive the item, it is as if he has gained three dollars.

In many settings with m items, $\mathbf{v}$ is a vector in $\mathbb{R}^m$, where v_j represents the buyer's value for item j. In more complex settings, $v(\cdot)$ might be a function of any set of items.

When there are multiple buyers, we denote agent i's type as v_i, the vector of all agents' types as $\mathbf{v}$, and the vector of all types aside from agent i as $\mathbf{v}_{-i}$. In general, $-i$ is used to denote all agents other than i.

(Direct) Mechanisms. For all settings in this thesis, the seller designs the mechanism. The mechanism consists of two rules: the allocation rule x and the payment rule p. In a direct mechanism, each buyer reports a bid b_i of his type. If his bid is *truthful*, then $b_i = v_i$. The allocation and payments are then computed on the reported bids of all buyers, $\mathbf{b}$. In mechanisms that are not direct, the buyers may take other actions and the allocation rules and payments may be computed based on these actions. However, we study direct mechanisms in this thesis.

Allocations. The allocation rule x is a function of the bids of all n buyers; $x_{ij}(b_1, \ldots, b_n)$ will denote the probability that item j is allocated to buyer i under bids $(b_1, \ldots, b_n)$. If $x_{ij}(b_1, \ldots, b_n) = 1$, then buyer i receives item j. If $x_{ij}(b_1, \ldots, b_n) = 0$, then buyers i does not receive item j. If $x_{ij}(b_1, \ldots, b_n) = \alpha \in (0, 1)$, then the mechanism flips a coin and with probability α, item j is allocated to buyer i.

Payments. The payment rule is also a function of the bids of all n buyers. If $p_i(b_1, \ldots, b_n) = 10$, then the mechanism charges buyer i \$10. We use the standard assumption of no-positive-transfers, that is, it is always the case that $p_i(\cdot) \geq 0$, for all i.

Utility. We assume that buyers have *quasi-linear* utility and are *risk-neutral*. Thus, if a buyer i has value v_i for an item that he receives with probability x and for which he is charged payment p (up front, independent of whether or not he successfully receives the item), then his utility is $xv - p$. His *allocated value* is the probability that he receives the item times his value for the time. *Quasi-linear* indicates that his utility is his allocated value minus his payment. *Risk-neutral* indicates that the contribution to his utility for getting an item with probability x is equal to x times his value, xv.

Consider a single-item setting with a buyer who has value v_i. If the buyer reports b_i and the other buyers report $\mathbf{b}_{-i}$, then we denote buyer i's utility as

$$u_i(b_i, \mathbf{b}_{-i} \mid v_i) := x_i(b_i, \mathbf{b}_{-i})v_i - p_i(b_i, \mathbf{b}_{-i}).$$

If buyer i reports truthfully, we shorthand this as

$$u_i(v_i, \mathbf{b}_{-i}) := u_i(v_i, \mathbf{b}_{-i} \mid v_i).$$

Environments. An auction environment describes the number of items for sale, the constraints on which buyers can be served simultaneously, and the sort of preferences that the buyers have. For each environment, we assume there are n buyers and m items.

In a *single-item* environment, the auctioneer sells $m = 1$ item. Each buyer's type v_i is a scalar describing his value for that item.

In a *single-parameter* environment, the auctioneer sells many items. However, each buyer's type v_i is still a scalar. Examples include:

- k-unit auctions, where each buyer has a value v_i and wants one item, but the seller has k identical items for sale.
- additive-up-to-k buyers, where m identical items are for sale, and each buyer has a a value of v_i for each item allocated to him up to k items, at which he gets no additional value.
- where m identical items are for sale and each buyer has a value v_i and wants one item, but there are feasibility constraints on which buyers the seller can simultaneously serve.
- where there are m distinct items, but there is a global belief about their ranking and quality, such as in position auctions, where if the buyer gets the j^{th} slot, his value is $\beta_j v_i$, where β_j is common and known to everyone.

A *single-dimensional* environment is one where the buyer's type is a scalar, such as single-item and single-parameter. A *multi-dimensional* is one where the buyer's type has multiple parameters. Typically, in a multi-dimensional environment, the buyer's type consists of at least m parameters, such as unit-demand and additive, which are defined below.

In a *unit-demand* environment, the auctioneer sells m items. Each buyer's type is an m-vector $\mathbf{v}_i \in \mathbb{R}^m$ where v_{ij} represents buyer i's value for item j. However, each buyer wants at most one item. That is, buyer i's value for receiving set S is $\max_{j \in S} v_{ij}$.

In an *additive* environment, the auctioneer sells m items. Each buyer's type is an m-vector $\mathbf{v}_i \in \mathbb{R}^m$ where v_{ij} represents buyer i's value for item j. Each buyer is additive over the items: for any set of items he receives, his value is the sum of the items contained in the set. That is, buyer i's value for receiving set S is $\sum_{j \in S} v_{ij}$.

Priors. In some scenarios, we assume that a buyer's type v is drawn from a prior distribution. We use F to denote the cumulative distribution function, and f to denote its density function. We also write $v \sim F$ to denote that v is drawn from the distribution (with CDF) F. When there are multiple buyers and multiple items, we use F to denote the joint distribution of all buyer types, and F_{ij} to denote the marginal distribution for buyer i's value for item j.

The Bayesian Setting. We say that we are in the *Bayesian setting* if the prior distribution of all buyers' values F is known to the auctioneer and all of the buyers.

Objectives. The mechanism designer aims to maximize some objective. The most common objectives are *social welfare*, also referred to as just "welfare" and also as "efficiency," and *revenue*.

Social welfare is thought of us "aggregate happiness," and is equal to aggregate buyer allocated value. Equivalently, it is the sum of *all* agents' utility, where the auctioneer herself counts as agent, and her utility is the revenue that she earns. Formally, for a mechanism defined by (x, p) and for reported bids $\mathbf{b}$,

$$\textsc{Welfare} := \sum_{i=1}^{n} x_i(\mathbf{b}) \cdot v_i$$

where $\cdot$ represents a dot product in the event that we are in an environment such as unit-demand or additive.

On the other hand, revenue is simply the sum of the buyers' payments, and this is thought of as "the seller's happiness." Since we cannot maximize each realization of payments, we instead maximize *expected revenue* over the prior distributions of the buyers. For truthfully reported values,

$$\textsc{Revenue} := \mathbb{E}_{\mathbf{v}\sim F}\left[\sum_{i=1}^{n} p_i(\mathbf{v})\right].$$

2.2 Incentives

In order to properly reason about agent's incentives and the guarantees that we get despite strategic behavior, we must define the following concepts.

Direct Revelation. The event where agents truthfully report their types to the mechanism.

Dominant Strategy Incentive Compatibility. A mechanism is *dominant-strategic incentive compatible* (DSIC) if, for every agent i, for any realization of agent i's type v_i, for any report of other agents' types $\mathbf{b}_{-i}$, buyer i maximizes his utility by reporting his true type v_i. That is,

$$u_i(v_i, b_{-i}|v_i) \geq u_i(b_i, b_{-i}|v_i) \quad \forall i, v_i, b_i, b_{-i}.$$

Ex-Post Incentive Compatibility. A mechanism is *ex-post incentive compatible* (EPIC) if, for every agent i, for any realization of agent types $\mathbf{v}$, and given that all other agents report their true values, buyer i maximizes his utility by reporting his true type v_i. That is, if $\mathbf{v}_{-i}$ are the true types of other agents,

$$u_i(v_i, \mathbf{v}_{-i}|v_i) \geq u_i(b_i, \mathbf{v}_{-i}|v_i) \quad \forall i, v_i, b_i, \mathbf{v}_{-i}.$$

Bayesian Incentive Compatibility. A mechanism is *Bayesian incentive compatible* (BIC) if, for every agent i, for any realization of agent i's type v_i, in expectation over the realization of the other agents' types $\mathbf{v}_{-i}$ and assuming that they report truthfully, buyer i maximizes his utility by reporting his true type v_i. That is, if $\mathbf{v}_{-i}$ are the true types of other agents, distributed according to CDF F_{-i}, then

$$\mathbb{E}_{\mathbf{v}_{-i}\sim F_{-i}}\left[u_i(v_i, \mathbf{v}_{-i}|v_i)\right] \geq \mathbb{E}_{\mathbf{v}_{-i}\sim F_{-i}}\left[u_i(b_i, \mathbf{v}_{-i}|v_i)\right] \quad \forall i, v_i, b_i.$$

Remark 1. *Observe that BIC implies EPIC which implies DSIC. Hence DSIC is the strongest form if IC, and BIC is the weakest, of these three forms.*

Individual Rationality. A mechanism is *ex-post individually rational* (IR) if, for every agent i, for any realization of agent types $\mathbf{v}$, every agent receives non-negative utility from participating in the mechanism *after the allocation has been determined*. That is, if $X_i(v_i, \mathbf{v}_{-i})$ is the realization of the (potentially randomized) allocation to buyer i and we overload notation and let $v_i(X_i(v_i, \mathbf{v}_{-i}))$ represent his valuation in any setting here, then

$$v_i(X_i(v_i, \mathbf{v}_{-i})) - p_i(v_i, \mathbf{v}_{-i}) \geq 0$$

A mechanism is *interim individually rational* if, for every agent i, for any realization of agent i's type v_i, every agent receives non-negative utility from participating in the mechanism *in expectation* over the realizations of the other agents' types and the randomness of the mechanism. That is,

$$\mathbb{E}_{\mathbf{v}_{-i}\sim F_{-i}}\left[u_i(v_i,\mathbf{v}_{-i})\right] \geq 0.$$

The VCG Mechanism. The Nobel-prize-winning VCG mechanism [Vickrey, 1961; Clarke, 1971; Groves, 1973] attains optimal social welfare and is DSIC. For any environment, let $x(\mathbf{v})$ denote the optimal allocation given n bidders with bids $\mathbf{v}$ and let $W^*(\mathbf{v})$ denote the social welfare of this allocation. Let $x_i(\mathbf{v})$ denote the set of items T_i^* that are allocated to buyer i in the optimal allocation $x(\mathbf{v})$. Let w_{i,T_i^*} denote the welfare that i contributes to W^* by receiving T_i^*. Then buyer i's payment is

$$p_i(\mathbf{v}) = W^*(\mathbf{v}_{-i}) - \left[W^*(\mathbf{v}) - w_{i,T_i^*}\right].$$

Observe that by definition of welfare, w_{i,T_i^*} is buyer i's allocated value under the optimal allocation, and thus his utility is $u_i(\mathbf{v}) = w_{i,T_i^*} - p_i(\mathbf{v}) = W^*(\mathbf{v}) - W^*(\mathbf{v}_{-i})$. Since buyer i has no influence over $W^*(\mathbf{v}_{-i})$ and the mechanism designer's goal is to maximize $W^*(\mathbf{v})$, then buyer i's incentives are aligned with the mechanism design, and his utility is best maximized by reporting his true value.

Revelation Principle. The revelation principle [Myerson, 1981] states that truthfulness essentially comes for free: if there is an optimal mechanism M (for whatever objective) that is not IC, then we design a new mechanism M' that is IC and achieves the same objective. The revelation principle depends on the definition of truthfulness. We sketch the idea for our above definitions of truthfulness, recalling what each definition means.

- DSIC: Truth-telling is a dominant strategy.
- EPIC: Truth-telling forms a Nash equilibrium.
- BIC: Truth-telling forms a Bayesian Nash equilibrium.

Suppose that each buyer i is incentivized to apply some strategy $s_i(\cdot)$ to his type (because it is his dominant or equilibrium strategy). That is, he reports $s_i(v_i)$ to the mechanism M. Then we can design a new mechanism M' that solicits bids from each buyer, and when receiving bid b_i from buyer i, applies $s_i(\cdot)$. Then buyer i is incentived to report v_i (as a dominant or equilibrium strategy), and M' achieves the same objective. See Figure 2.1.

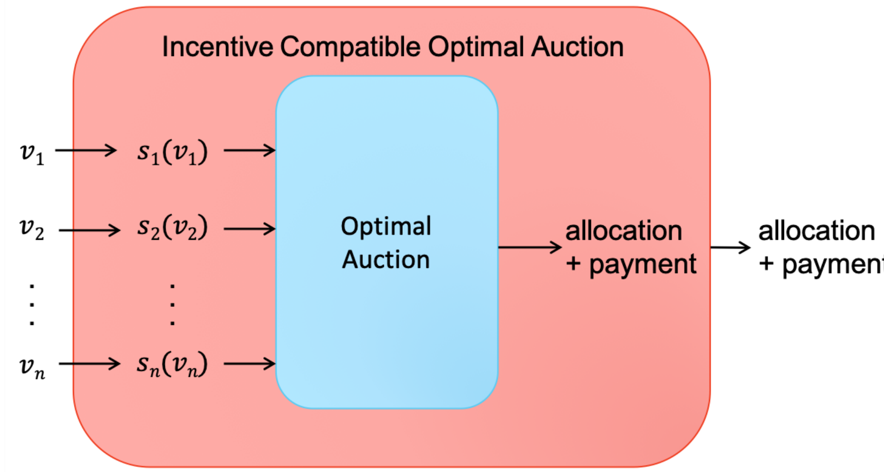


Figure 2.1: Depiction of the revelation principle.

Monotonicity and Payment Identity. In the single-parameter environment, Myerson [1981] proves that DSIC is equivalent to the allocation rule x satisfying *monotonicity* in each agent's bid and the *payment identity*. Monotonicity in the allocation rule is when the probability that the item is allocated to buyer i increases as buyer i's bid b_i increases: $x_i(b_i, \mathbf{b}_{-i}) \geq x_i(b'_i, \mathbf{b}_{-i})$ when $b_i > b'_i$. The payment identity is as follows:

$$p_i(b_i, \mathbf{b}_{-i}) := b_i \cdot x_i(b_i, \mathbf{b}_{-i}) - \int_0^{b_i} x_i(z, \mathbf{b}_{-i}) dz.$$

In the following derivation, we omit the argument of the other bids $\mathbf{b}_{-i}$ in the allocation and payment function, as they remain fixed throughout the entire argument.

From DSIC, we have that

$$vx_i(v) - p_i(v) \geq vx_i(v-\varepsilon) - p_i(v-\varepsilon) \quad \text{and} \quad (v-\varepsilon)x_i(v-\varepsilon) - p_i(v-\varepsilon) \geq (v-\varepsilon)x_i(v) - p_i(v)$$

then

$$v\left[x_i(v) - x_i(v-\varepsilon)\right] \geq p_i(v) - p_i(v-\varepsilon) \geq (v-\varepsilon)\left[x_i(v) - x_i(v-\varepsilon)\right]$$

and by taking $\lim_{\varepsilon \to 0}$, we get that

$$vx_i'(v) = p_i'(v)$$

where the $'$ denotes the derivative with respect to v. Integration then gives that

$$p_i(v) = \int_0^v p_i'(z)dz = \int_0^v zx_i'(z)dz = vx_i(v) - \int_0^v x_i(z)dz$$

where the last equality follows from integration by parts.

Utility is Area Under the Allocation. Given the payment identity, in the single-parameter setting, then in a DSIC mechanism, agent i's utility is

$$u_i(v_i, \mathbf{v}_{-i}) = v_i \cdot x_i(\mathbf{v}) - p_i(\mathbf{v}) = \int_0^v x_i(z, \mathbf{v}_{-i})dz.$$

That is, holding the values of other agents $\mathbf{v}_{-i}$ fixed, the utility of agent i for having type v_i is the *area under the allocation curve* x up until v_i.

Taxation Principle. The taxation principle states that any direct revelation mechanism is equivalent to offering a menu of lotteries. A lottery consists of a price and a probability of allocation for every item. In essence, every report to the mechanism corresponds to some menu option. The buyer could purchase any menu option, but since truth-telling maximizes his utility, then he prefers to purchase menu option that corresponds to reporting his true type. Thus, any mechanism is equivalent to offering a menu of (price, allocation) pairs.

Menu Complexity. For a single buyer, one metric of complexity of a mechanism is the number of menu options that are offered in the menu form of the mechanism, excluding the option $(0, 0)$ (which must be included in all menus in order to be individually rational).

2.3 Revenue Maximization.

The following concepts are defined according to Myerson's single-parameter revenue-optimal theory. All of the concepts will be defined in terms of n identical buyers and 1 item, but can be generalized to n non-identical buyers.

Revenue Curves. The *revenue curve* for an item with CDF F is a function R that maps a value v to the revenue obtained by posting a price of v for that item when buyer values are drawn from the distribution F. Formally, $R(v) := v \cdot [1 - F(v)]$.

Reserve Prices. The *monopoly reserve price* for an item where buyer values are drawn from F and has revenue curve $R(\cdot)$ is $r \in \mathrm{argmax}_p R(p)$.

Quantile Space. We can define the quantile of a value $q_F(v)$ as probability that a buyer has value at least v: $q_F(v) = 1 - F(v)$. Then, for a quantile $q \in [0, 1]$, we define the value corresponding to q as the unique value

$$v_F(q) = F^{-1}(1 - q) = \sup\{v \mid 1 - F(v) \leq q\}.$$

We can redefine the revenue curve in quantile space as $R(q) = q \cdot v_F(q)$.

Virtual Values. Myerson's virtual valuation function $\varphi(\cdot)$ is defined so that $\varphi(v) := v - \frac{1-F(v)}{f(v)}$. This is the negative derivative of the revenue curve in quantile space: $-\frac{d}{dq}R(q) = \varphi(q) = \varphi(v)$ where $v = v_F(q)$.

Regularity. We say that a distribution F satisfies *regularity* if the virtual values $\varphi(\cdot)$ for this distribution are monotone non-decreasing, or equivalently, if the revenue curve in quantile space $R(q)$ is concave.

Ironing. When a distribution does not satisfy the regularity assumption, we instead need to *iron* the virtual values.

The *ironed revenue curve* denoted $\hat{R}(\cdot)$ for a revenue curve $R(\cdot)$ is the least concave upper bound on the revenue curve $R(\cdot)$ *in quantile space*. A point v is *ironed* if $\hat{R}(q_F(v)) \neq R(q_F(v))$. We say that $[a, b]$ is an *ironed interval* if $\hat{R}(q_F(a)) = R(q_F(a))$,

$\hat{R}(q_F(b)) = R(q_F(b))$, and $\hat{R}(q_F(v)) \neq R(q_F(v))$ for all $v \in (a, b)$, where if $v \in (a, b)$, then a and b are the lower and upper endpoint of the ironed interval, respectively.

The *ironed virtual values* denoted $\hat{\varphi}(\cdot)$ for original virtual values $\varphi(\cdot)$ are the negative derivative of the ironed revenue curve in quantile space. That is, $\hat{\varphi}(q) = -\frac{d}{dq}\hat{R}(q)$. The ironed virtual values in value space are simply converted to value space: $\hat{\varphi}(v) = \hat{\varphi}(q_F(v))$.

Because $\hat{\varphi}(\cdot)$ is the derivative of the least concave upper bound, then throughout ironed intervals, the slope will be linear, and thus the derivative of $\hat{R}(q)$ and also the ironed virtual values will be constant.

Virtual Welfare. Myerson [1981] proved that expected revenue is equal to expected virtual welfare:

$$\textsc{Revenue} = \mathbb{E}_{\mathbf{v}\sim F}\left[\sum_i x_i(\mathbf{v})\varphi_i(v_i)\right].$$

This is the "social welfare" for an allocation computed with the virtual values instead of with the buyers' actual values. Then if a mechanism designer wishes to maximize expected revenue, she can instead maximize expected virtual welfare, and can express the expected revenue objective without using payment variables. We re-derive this here. The second equality uses the payment identity, and the rest is just algebra.

$$\begin{aligned}
\mathbb{E}_{v_i\sim F_i}\left[p_i(\mathbf{v})\right] &= \int_0^\infty f_i(v_i)p_i(\mathbf{v})dv_i \\
&= \int_0^\infty f_i(v_i)\left[v_i x_i(\mathbf{v}) - \int_0^{v_i} x_i(z, \mathbf{v}_{-i})dz\right]dv_i \\
&= \int_0^\infty f_i(v_i)v_i x_i(\mathbf{v}) - \int_0^\infty x(v_i, \mathbf{v}_{-i})\int_{v_i}^\infty f_i(w)dwdv_i \\
&= \int_0^\infty f_i(v_i)v_i x_i(\mathbf{v}) - \int_0^\infty x(\mathbf{v})\left[1 - F_i(v_i)\right] \\
&= \int_0^\infty f_i(v_i)x_i(\mathbf{v})\left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}\right] \\
&= \mathbb{E}_{v_i\sim F_i}\left[x_i(\mathbf{v})\varphi_i(v_i)\right]
\end{aligned}$$

Then

$$\textsc{Revenue} = \mathbb{E}_{\mathbf{v}\sim F}\left[\sum_i p_i(\mathbf{v})\right] = \mathbb{E}_{\mathbf{v}\sim F}\left[\sum_i x_i(\mathbf{v})\varphi_i(v_i)\right].$$

2.4 Lagrangian Duality

In this section, we provide the basics surrounding formulating a partial Lagrangian primal and taking its dual, and understanding the properties of duality. We use these properties in our duality techniques in chapters 3-5.

We begin with a standard maximization problem subject to constraints, which we call the full primal. The set $\mathcal{P}$ here denotes feasibility constraints, while x represents whatever our primal variables are.

Full primal:

$$\begin{aligned} \max \quad & f(x) & \\ \text{s.t.} \quad & Ax \le b & \text{(dual variable } \lambda\text{)} \\ & x \in \mathcal{P} & \end{aligned}$$

We denote the optimal solution to the full primal as x^*; that is, $x^* \in \text{argmax}_{Ax\le b, x\in\mathcal{P}} f(x)$.

We now form the partial Lagrangian primal by using the Lagrangian multiplier λ_i for each constraint of the form $(Ax)_i \le b_i$ and moving it into the objective, where we now minimize over the multipliers λ. We leave all of the feasibility constraints as is, and define $\mathcal{L}(x;\lambda)$ as this new objective.

Lagrangian Primal:

$$\max_{x\in\mathcal{P}} \min_{\lambda\ge 0} \quad \mathcal{L}(x;\lambda) \quad = \quad \max_{x\in\mathcal{P}} \min_{\lambda\ge 0} \quad f(x) + \lambda^T(b - Ax)$$

By reversing the order of the max and the min, we obtain the dual minimization problem. We notate this dual problem as $D(\lambda)$.

Lagrangian Dual:

$$\min_{\lambda\ge 0} \quad D(\lambda) \quad = \quad \min_{\lambda\ge 0} \max_{x\in\mathcal{P}} \quad f(x) + \lambda^T(b - Ax)$$

We denote the optimal dual solution as $\lambda^* \in \text{argmin}_{\lambda \geq 0} D(\lambda)$.

We say that x, λ satisfy complementary slackness if $\lambda_i \geq 0 \implies b_i - (Ax)_i = 0$.

Relaxation. First, we observe that the (partial) Lagrangian Primal is in fact a relaxation of the full primal. For any feasible x, λ—that is, $Ax \leq b, x \in \mathcal{P}$, and $\lambda \geq 0$—then $f(x) \leq \mathcal{L}(x; \lambda)$.

Weak Duality. The value of the full primal is always upper-bounded by the value of the dual problem. Specifically, the value of the full primal is at most $f(x^*)$ by definition, and any feasible dual solution must satisfy $\lambda \geq 0$, so the dual objective is larger: $f(x^*) \leq D(\lambda)$.

Proof.

$$\begin{aligned} f(x^*) &\leq f(x^*) + \lambda^T(b - Ax^*) && \lambda \geq 0, Ax^* \leq b \\ &\leq \max_{x \in \mathcal{P}} f(x) + \lambda^T(b - Ax) && x^* \in \mathcal{P} \\ &= D(\lambda) \end{aligned}$$

□

Strong Duality. Strong duality implies that the value of the full primal is equal to the value of the Lagrangian primal, and this is equal to the value of the Lagrangrian dual, when they are all at their optimal solutions. However, strong duality is not a given. We see below that if strong duality holds, there must exist a pair of primal, dual solutions that are optimal. Further, if there exist an optimal pair, then strong duality must hold. Either condition is sufficient to show the other exists.

An Optimal Pair implies Strong Duality. For any choice of dual variables $\hat{\lambda}$, if there exists $\hat{x}$ that forms an optimal pair with $\hat{\lambda}$, that is, $\hat{x}$ such that:

1. $\hat{x} \in \text{argmax}_{x \in \mathcal{P}} \mathcal{L}(x; \hat{\lambda})$ ($\hat{x}$ is optimal)
2. $A\hat{x} \leq b$ ($\hat{x}$ satisfies the Lagrangified constraints)
3. $\hat{x}, \hat{\lambda}$ satisfy complementary slackness

then strong duality holds, that is, $D(\hat{\lambda}) = f(x^*)$.

Proof.

$$\begin{aligned} D(\hat{\lambda}) &= \max_{x \in \mathcal{P}} \mathcal{L}(x, \hat{\lambda}) \\ &= f(\hat{x}) + \hat{\lambda}^*(b - A\hat{x}) && \text{by (1)} \\ &= f(\hat{x}) && \text{by (3)} \\ &\leq f(x^*) && \text{by (2), } x \in \mathcal{P} \end{aligned}$$

□

Strong Duality implies an Optimal Pair. If strong duality holds, that is, $\min_{\lambda \geq 0} D(\lambda) = f(x^*)$, then there exists $\hat{x}$ such that

1. $\hat{x} \in \operatorname{argmax}_x \mathcal{L}(x; \lambda^*)$
2. $A\hat{x} \leq b$
3. $\hat{x}, \lambda^*$ satisfy complementary slackness
4. $f(\hat{x}) = f(x^*)$.

Proof. From weak duality, we know that

$$\min_{\lambda \geq 0} D(\lambda) \quad = \quad D(\lambda^*) \quad \geq \quad \mathcal{L}(x^*, \lambda^*) \quad \geq \quad f(x^*).$$

These inequalities must all hold with equality for the premise to hold. The first inequality's tightness implies condition (1), and the second inequality's tightness implies condition (3). Condition (2) is true by the definition of x^*. □

For further background on Lagrangian duality, see [Rockafellar, 1974].

Part I

Interdimensional Mechanism Design

3 The FedEx Problem

"Remember that Time is Money"

— Benjamin Franklin in *Advice to a Young Tradesman* (1748)

3.1 Introduction

Consider the pricing problem faced by FedEx. Each of their customers has a deadline d by which he needs his package to arrive, and a value v for receiving the package by the deadline. The customer's utility for getting his package shipped by day i at a price of p is $v - p$ if $i \leq d$ (i.e., it is received by his deadline) and $-p$ otherwise. Of course, a customer's (v, d) pair is the private information of the customer. We study the Bayesian setting, where this pair (v, d) is drawn from a prior distribution known to FedEx, and address the question of optimal (revenue maximizing) mechanism design. Note that the prior distribution may be arbitrarily correlated.

Suppose that FedEx offers a discrete set of shipping options (1-day, 2-day, 3-day, up to m-day shipping). The prior that FedEx has on its customer's needs is given by a probability distribution $(q_1, q_2, \ldots, q_m)$, where q_i is the probability that the customer has a deadline i days from now, and a set of marginal value distributions, where F_i, for $1 \leq i \leq m$, is the distribution of values given that the customer's deadline is i.

We consider the single agent problem in this chapter, or equivalently, the setting where FedEx sells to identically drawn buyers and has constraints on the number of buyers it can supply. We obtain a closed form, clean, and optimal auction for this setting. Our work adds to the relatively short list of multi-parameter settings where a closed form solution is known. (See related work below for more on this.)

The pricing problem we consider is extremely natural and arises in numerous scenarios, whether it is Amazon.com providing shipping options, Internet Service Providers offering bandwidth plans, or a myriad of other settings in which a seller can price discriminate or otherwise segment her market by delaying service, or providing lower quality/cheaper versions of a product. In particular, this setting is relevant whenever a customer has a value and a sensitivity to time or some other

This chapter is based on joint work with Amos Fiat, Anna Karlin, and Elias Koutsoupias in a paper by the same title which appeared at EC 2016 [FGKK '16].

feature of service. A "deadline" represents the base level of need, imposed on a buyer by outside circumstances, whereas a valuation represents the buyer's own willingness to pay. It is important to understand how buyer deadline constraints impact the design of auctions and what leverage they give to the auctioneer to extract more revenue.

3.1.1 Related Work

The FedEx setting is a variant of price discrimination in which the customers are grouped by their deadline. Price discrimination offers different prices to users with the goal of improving revenue [Bergemann, Brooks, and Morris, 2015]. Alternatively one can view the FedEx problem as a multi-dimensional optimal auction problem. There are two ways to express the FedEx problem in this way. First, as a 2-dimensional (value $\times$ deadline) problem of arbitrary joint distribution in which the second variable takes only integer values in a bounded interval. Alternatively, as a very special case of the m-dimensional unit-demand problem with correlated values (the customer buys a shipping option among the m choices)—his value for the first d options is v, and for the last $m - d$ is 0.

There is an extensive body of literature on optimal auction design. The seminal work of Myerson [1981] has completely settled the case of selling a single item to multiple bidders and extends directly to the more general framework of single-parameter settings. Note that Riley and Zeckhauser [1983] also prove that the optimal single-parameter mechanism is deterministic, and that Bulow and Roberts [1989] are responsible for the interpretation of virtual values as a marginal contribution to revenue.

The most complicated part of Myerson's solution is his handling of distributions that are not regular by "ironing" them, that is, by replacing the revenue curves by their upper concave envelope. Myerson's ironing is done in quantile space. In this work, we also need to iron the revenue curves, but we need to do this in value space.

Extending Myerson's solution to the multi-dimensional case has been one of the most important open problems in Microeconomics. For the case of unit-demand agents, a beautiful sequence of papers [Chawla, Hartline, and Kleinberg, 2007; Briest et al., 2015; Chawla, Malec, and Sivan, 2015; Chawla et al., 2010; Alaei, 2011; Cai and Daskalakis, 2011] showed how to obtain approximately optimal auctions. For the case of finite type spaces, [Cai, Daskalakis, and Weinberg, 2012, 2013a,b] are able to use linear and convex programming techniques to formulate and solve the optimal auction problem. This gives a black-box reduction from mechanism to algorithm design that yields a PTAS for revenue maximization in unit-demand settings. For the case of additive agents, additional recent breakthroughs [Hart and Nisan, 2017; Li and Yao, 2013; Babaioff et al., 2014; Yao, 2015; Cai et al., 2016; Cai and Huang, 2013] have also resulted in approximately optimal mechanisms. See Section 5.1.1 for more details on this line of work.

But if we insist on optimal auctions for continuous probability distributions, no general solution is known for the multi-dimensional case—even for the two-dimensional single-bidder case—and it is very possible that no such simple solution exists for the general case. One of the reasons that the multi-dimensional case is so complex is that optimal auctions are not necessarily deterministic [Pavlov, 2011; Thanassoulis, 2004; Briest et al., 2015; Hart and Reny, 2012; Hart and Nisan, 2013; Manelli and Vincent, 2006; Pycia, 2006; Daskalakis et al., 2013]. The optimal auction for the FedEx problem also turns out to be randomized with exponentially many different price levels in the worst case [Saxena et al., 2018].

There are some relevant results that solve special cases of the two-parameter setting. One of the earliest works is by Laffont, Maskin, and Rochet [1987] who study a distant variant of the FedEx problem. In their problem, the bidder has two parameters, a and b, each uniformly distributed on $[0, 1]$, and the bidder's utility function is very specifically the quadratic function $ax - \frac{1}{2}(b + 1)x^2 - p$. Here, x is a single-dimensional allocation variable and p is the payment variable. The idea is that both the slope and the intercept of the buyer's demand curve are unknown to the seller. To solve this problem, Laffont et al. come up with a change of variable technique to use only one variable when solving for the allocation in both parameters. By solving the optimization problem and the resulting integration by hand with this technique, they provide a highly non-trivial closed-form allocation rule, demonstrating that even the simplest independent two-parameter settings are far more difficult than single-parameter settings.

McAfee and McMillan [1988] study a generalization of this problem. First, they characterize incentive-compatibility precisely in direct, deterministic, and differentiable mechanisms. Then, they reference a notion of "single-crossing" which says that the marginal rate of substitution[1], must be monotone increasing in the

[1]This is equal to the derivative of utility with respect to the allocation divided by the negative

buyer's type. McAfee and McMillan generalize this condition to multi-parameter settings, and then extend the analysis of Laffont et al. (using the same change of variables technique) to any number of variables if they satisfy generalized single-crossing and other small conditions. However, their analysis only applies when the optimal mechanism is deterministic. Finally, McAfee and McMillan also consider the setting where a buyer has independent valuations for m heterogenous items. They prove that under a "regularity" condition[2], for $m = 2$, the optimal mechanism is deterministic, and they further reason from prior results that the optimal mechanism would set a price for each item individually as well as the grand bundle.

These initial results were followed by more general results. In particular, Haghpanah and Hartline [2015] consider the problem of selling a product with multiple quality levels to unit-demand bidders. The mechanism they consider is selling only the highest quality product at a posted price. (In the FedEx problem, this corresponds to having a single price for every shipping option.) When the buyers' value distributions have a specific type of positive correlation, then this mechanism is optimal, because the high-valued customers are less quality-sensitive, and thus will not pay a premium for a different outcome. Haghpanah and Hartline solve for when this mechanism is a point-wise virtual value maximizer, with expected revenue equal to virtual welfare, and then solve for the paths of tight IC constraints to integrate over, effectively reverse engineering the virtual value functions. This approach also corresponds to proving when bundling is optimal in particular additive settings. Their work generalizes results from Armstrong [1996].

Daskalakis et al. [2017] establish a duality framework where the primal is expressed in terms of utility and a transformed measure μ of the buyer distributions, and the dual is an optimal transport problem. The dual variables are a measure μ' that stochastically dominate the primal measure μ, and the objective is the distance between the positive part of the measure μ'_+ and the negative part of the measure μ'_-. First, Daskalakis et al. establish that strong duality holds in their duality framework, so the dual can be used to solve for or certify the optimal primal. Then, using their duality framework, for any mechanism with a finite menu size (number of outcomes), they give a characterization in terms of the stochastic dominance conditions on the measure μ. One application of this result proves that for m items distributed

derivative of utility with respect to payment. This also the negative shadow price.

[2]They assume that $t \cdot f'(t) + (m+1)f(t) \geq 0$. Note that this is *not* Myerson's regularity condition, but would be if the $(m+1)$ were replaced by 1.

i.i.d. on $[c, c+1]$ for large enough c, grand bundling is optimal.

Our approach is based on a duality framework. Two such frameworks have been proposed. The first framework by Daskalakis et al. [2017] expresses the primal in terms of utility and a transformed measure μ of the buyer distributions, and the dual is an optimal transport problem. The dual variables are a measure μ' that stochastically dominate the primal measure μ, and the objective is the distance between the positive part of the measure μ'_+ and the negative part of the measure μ'_-. First, Daskalakis et al. establish that strong duality holds in their duality framework, so the dual can be used to solve for or certify the optimal primal. Then, using their duality framework, for any mechanism with a finite menu size (number of outcomes), they give a characterization in terms of the stochastic dominance conditions on the measure μ. One application of this result proves that for $m-1$ items distributed i.i.d. on $[c, c+1]$ for large enough c, grand bundling is optimal.

The second framework is by Giannakopoulos and Koutsoupias [2014], which is based on expressing the revenue maximization problem as an optimization problem in terms of utility functions and their partial derivatives. They find primal and dual variables that are both feasible and also satisfy complementary slackness, and thus via weak duality and complementarity, have equal objectives. Using their framework, they prove that the Straight-Jacket-Auction[3] is optimal for an additive bidder whose item valuations are i.i.d. from $U[0,1]$ for up to six items. In [Giannakopoulos and Koutsoupias, 2015], the authors subsequently use their framework to give closed-form optimal allocation and payment rules for several independent non-identical two-item problems (where the distributions are from monomial or exponential families over $[0, H]$), and the mechanisms are no longer deterministic. The duality framework of Giannakopoulos and Koutsoupias is a fairly general approach, but their applications still require fairly strong assumptions about the distributions in order to make progress on characterizing optimal auctions. Our solution of the FedEx problem follows this latter duality framework.

For much more on both exact and approximate optimal mechanism design, see [Daskalakis, 2015; Chawla and Sivan, 2014; Roughgarden, 2015; Hartline, 2013; Cai, Daskalakis, and Weinberg, 2011]. For background on duality in infinite linear and

[3]The Straight-Jacket-Auction for m items is a deterministic mechanism where a price is set for every bundle size, bound by sale probabilities. The price for the bundle of size r, $p_r^{(m)}$, is such that, given prices $p_1^{(m)}$ through $p_{r-1}^{(m)}$ already set, a buyer with value $v \sim [0,1]^r \times 0^{m-r}$ will not buy any bundle of size r or smaller with probability $1 - \frac{r}{m+1}$.

convex programs, see e.g., [Anderson and Nash, 1987; Luenberger, 1997].

3.1.2 Our Contribution

Our result is one of the first explicit closed-form generalizations of [Myerson, 1981] to multi-parameter settings with arbitrary (joint) distributions, and contributes to recent breakthroughs in this space. We use a duality framework where we prove optimality by finding primal and dual solutions that satisfy sufficient conditions. The optimal primal and dual variables have an interesting inductive structure, and the allocation rule is potentially randomized over at most 2^{i-1} prices on i-day. Our approach strengthens the emerging understanding that duality is useful for determining the structure of the optimal auction in non-trivial settings, in addition to its use in analyzing the auction.

In Myerson's setting, the "ironing" of revenue curves and virtual valuations to determine the optimal auction is required to enforce incentive compatibility constraints among multiple bidders. In our setting, we need a form of ironing even for one bidder in order to enforce incentive compatibility constraints among the multiple options. This work also suggests that ironing is one of the biggest hurdles in extending Myerson to more general settings.

3.2 Preliminaries

As discussed above, the type of a customer is a (value, deadline) pair. An auction takes as input a reported type $t = (v, d)$ and determines the shipping date in $\{1, \ldots, m\}$ and the price. We denote by $a_i(v)$ the probability that the package is shipped by day i when the agent reports (v, i), and by $p_i(v)$ the corresponding expected payment (the expectation is taken over the randomness in the mechanism).

Our goal is to design an optimal mechanism for this setting. By the revelation principle, we can restrict our attention to incentive compatible mechanisms. In this setting, when an agent with type (v, i) reports a type of (v', i'), he has utility

$$u(v', i' \mid v, i) = \begin{cases} va_{i'}(v') - p_{i'}(v') & \text{if } i' \leq i \\ -p_{i'}(v') & \text{otherwise.} \end{cases}$$

The incentive compatibility requirement is that

$$u(v,i) \geq u(v',i' \mid v,i) \qquad \forall v',i'. \tag{3.1}$$

We also require individual rationality, i.e., $u(v,i) \geq 0$ for all (v,i). Without loss of generality, $a_i(v)$ is the probability that the package is delivered *on day* i, since any incentive compatible mechanism which delivers a package early can be converted to one that always delivers on the deadline, while retaining incentive compatibility and without losing any revenue.

For each fixed i, this implies the standard (single parameter) constraints [Myerson, 1981], namely

$$\forall i,\ a_i(v) \text{ is monotone weakly increasing and in } [0,1]; \tag{3.2}$$

$$\forall i,\ p_i(v) = va_i(v) - \int_0^v a_i(x)dx \quad \text{and hence} \quad u(v,i) = \int_0^v a_i(x)dx. \tag{3.3}$$

Clearly no agent would ever report $i' > i$, as this would result in non-positive utility. However, we do need to make sure that the agent has no incentive to report an earlier deadline, and hence another IC constraint is that for all $2 \leq i \leq m$:

$$u(v,i-1|v,i) \leq u(v,i) \tag{3.4}$$

which is equivalent to

$$\int_0^v a_{i-1}(x)dx \leq \int_0^v a_i(x)dx \qquad \forall i \text{ s.t. } 1 < i \leq m. \tag{3.5}$$

We sometimes refer to this as the inter-day IC constraint. Since $a_i(v)$ is the probability of allocation of i-day shipping given report (v,i), constraints (3.2), (3.3) and (3.5) are necessary and sufficient, by transitivity, to ensure that

$$u(v,i) \geq u(v',i'|v,i)$$

for *all* possible misreports (v',i').

The prior. We assume that the agent's (value, deadline) comes from a known joint prior distribution F. Let q_i be the probability that the customer has a deadline

$i \in \{1, \ldots, m\}$, that is,

$$q_i = \Pr_{(v,d)\sim F}[d = i]$$

and let $F_i(\cdot)$ be the marginal distribution function of values for bidders with deadline i. That is,

$$F_i(x) = \Pr_{(v,d)\sim F}[v \le x \mid d = i].$$

We assume that F_i is atomless and strictly increasing, with density function defined on $[0, H]$. Let $f_i(v)$ be the derivative of $F_i(v)$.

The objective. Let $\varphi_i(v) = v - \frac{1-F_i(v)}{f_i(v)}$ be the virtual value function for v drawn from distribution F_i. Applying the Myerson payment identity (3.3) implies that the expected payment of a customer with deadline i is

$$\mathbb{E}_{v\sim F_i}[p_i(v)] = \mathbb{E}_{v\sim F_i}[\varphi_i(v)a_i(v)].$$

Thus, we wish to choose monotone allocation rules $a_i(v)$ for days $1 \le i \le m$, so as to maximize

$$\mathbb{E}_{(v,i)\sim F}[p_i(v)] = \sum_{i=1}^{m} q_i \mathbb{E}_{v\sim F_i}[p_i(v)] = \sum_{i=1}^{m} q_i \mathbb{E}_{v\sim F_i}[\varphi_i(v)a_i(v)] = \sum_{i=1}^{m} q_i \int_0^H \varphi_i(v) f_i(v) a_i(v) dv,$$

subject to (3.2) and (3.5).

A trivial case and discussion. If we knew that the customer that would arrive would have deadline i and we could thus condition on this event, ensuring that his value is drawn from the marginal distribution F_i, then the optimal pricing would be trivial, as this is a single-agent, single-item auction. In this case, the optimal mechanism for such a customer is to set the price for service by day i to the reserve price r_i for his prior. If we just had a number of single-dimensional problems, one for each deadline, we would want to set a price of r_i for each i-day shipping option. If it is the case that $r_i \ge r_{i+1}$ for each i, then the entire Fedex problem is trivial, since setting r_i as the price for i-day shipping satisfies all of the IC constraints, and this pointwise optimizes each conditional distribution.

Note that even should the marginal distribution for buyer values for 1-day shipping stochastically dominate the marginal distribution for 2-day shipping and so on, the later shipping options may still have higher reserve prices. For example, if F_2 is uniform over the set $\{1, 10\}$, the reserve is 10. If F_1 is uniform over the set $\{9, 10\}$, the reserve is 9. Hence, we do not make the assumption that the reserve prices are weakly decreasing with the deadline.

In fact, we do not make the assumption of stochastic domination either in order to be fully general. The prior F captures the result of random draws from a population consisting of a mixture of different types. Obviously any particular individual with deadline i is at least as happy with day $i - 1$ service as with day i service, but two random individuals may have completely uncorrelated needs. To give an example, an individual ordering a last minute birthday present may have a lower value than an individual scheduling the delivery of surgical equipment that is needed to perform open heart surgery in three weeks time. In fact, for more valuable packages, one could imagine that people take the time to plan ahead.

Another factor has to do with costs. It is likely that the cost that FedEx incurs for sending a package within i days is higher than the cost FedEx incurs for sending a package within $i' > i$ days, since in the latter case, for example, FedEx has more flexibility about which of many planes/trucks to put the package on, and even may be able to reduce the total number of plane/truck trips to a particular destination given this flexibility. More generally, in other applications of this problem, the cost of providing lower quality service is lower than the cost of providing higher quality service. Thus, even if reserve prices tend to decrease with i, all bets are off once we consider a customer's value for deadline i conditioned on that value being above the expected cost to FedEx of shipping a package by deadline i for each i.

In this chapter, we are not explicitly modeling the costs that FedEx incurs, the optimization problems that it faces, the online nature of the problem, or any limits on FedEx's ability to ship packages. These are interesting problems for future research. The discussion in the preceding few paragraphs is here merely to explain why the problem remains interesting and relevant even in the with distributions that do not have decreasing reserve prices in the deadlines. Further, note that in Figure 3.1, FedEx actually did post a larger price for a later shipping option, implying that they estimate the underlying distributions to have increasing reserve prices.[4]

[4]Of course, these prices are not incentive compatible, and the author that purchased shipping when presented with these prices did in fact misreport her deadline.

3. Rates and Transit Times Help

Amounts are shown in USD

Select	Delivery Date/Time	Service	Rates
	In the shortest time possible. Call 1·800·Go·FedEx for availability and rate.	FedEx International Next Flight®	
○	Sun Feb 26, 2017 by 6:00 PM	FedEx International Priority®	70.02
○	Wed Mar 1, 2017 by 6:00 PM	FedEx International Economy®	97.03

Figure 3.1: FedEx posted a higher price posted for a later shipping option, implying the underlying distributions do not have decreasing reserve prices in the real world. Note also that these prices are not incentive compatible.

3.3 Warm-up: The case of $m = 2$

Suppose that the customer might have a deadline of either one day or two days from now. By the taxation principle, the optimal mechanism is a menu, and in this setting consists of a (potentially randomized) price p_i for having the package delivered i days from now.

Let $R_i(v)$ be the i-day *revenue curve*, that is, $R_i(v) := v \cdot [1 - F_i(v)]$. Let $r_i := \text{argmax}_v R_i(v)$ be the price at which expected revenue from a bidder with value drawn from F_i is maximized, and let $R_i^* := R_i(r_i)$ denote this maximum expected revenue. Since R_i^* is the optimal expected revenue from the agent [Myerson, 1981], conditioned on having a deadline of i, then $q_1 R_1^* + q_2 R_2^*$ is an upper bound on the optimal expected revenue for the two-day FedEx problem. If $r_1 \geq r_2$, then this optimum is indeed achievable by an IC mechanism: just set the 1-day shipping price p_1 at r_1 and the 2-day shipping price p_2 at r_2.

But what if $r_2 > r_1$? In this case, the inter-day IC constraint (3.5) is violated by this pricing (a customer with $i = 2$ will pretend his deadline is $i = 1$).

Attempt #1: One alternative is to consider the optimal single price mechanism (i.e., $p_1 = p_2 = p$). In this case, the optimal choice is clear:

$$p := \text{argmax}_v \left[q_1 R_1(v) + q_2 R_2(v) \right], \tag{3.6}$$

i.e., set the price that maximizes the combined revenue from both days. There are cases where this is optimal, e.g., if both F_1 and F_2 are regular. A proof is given in Subsection 3.4.2.

Attempt #2: Another auction that retains incentive compatibility, and, in some cases, improves performance is to set the 1-day price p_1 at p and the 2-day price at

$$p_2 := \text{argmax}_{v \le p} R_2(v). \tag{3.7}$$

However, even if we fix $p_1 = p$, further optimization may be possible if F_2 is not regular.

Attempt #3: Consider the concave hull of $R_2(\cdot)$, i.e., the ironed revenue curve. If $R_2(v)$ is maximized at $r_2 > p$ and $R_2(\cdot)$ is ironed at p, then offering a lottery on 2-day with an expected price of p yields higher expected revenue than offering any deterministic 2-day price of p_2. As we shall see, for this case, this solution is actually optimal. (See Figure 3.2.)

However, if $p > r_2$, (which is possible if F_1 and F_2 are not regular, even if $r_1 < r_2$), then we will see that the optimal 1-day price is indeed higher than r_2, but not necessarily equal to p.

Attempt #4: If $p > r_2$, set the 1-day price at

$$p_1 := \text{argmax}_{v \ge r_2} R_1(v).$$

This should make sense: if we're going to set a 1-day price above r_2, we may as well set the 2-day price at r_2, but in that case, the 2-day curve should not influence the pricing for 1-day (except to set a lower bound for it).

Admittedly, this sounds like a tedious case analysis, and extending this reasoning to three or more days gets much worse. Happily, though, there is a nice, and relatively simple way to put all the above elements together to describe the solution, and then, as we shall see in Section 3.5, prove its optimality via a clean duality proof.

A solution for $m = 2$. Define $\hat{R}(\cdot)$ to be the concave (ironed) revenue curve corresponding to revenue curve $R(\cdot)$. We define the following combined revenue curve, depicted in Figure 3.3. Let

$$R_{12}(v) := \begin{cases} q_1 R_1(v) + q_2 \hat{R}_2(v) & v \le r_2 \\ q_1 R_1(v) + q_2 R_2(r_2) & v > r_2. \end{cases} \tag{3.8}$$

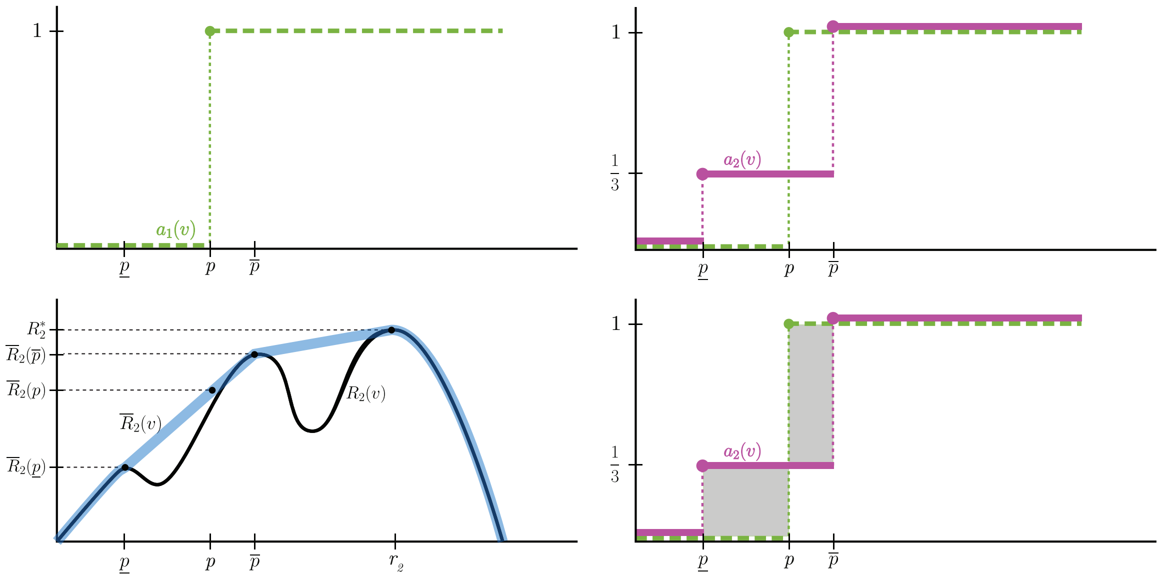


Figure 3.2: A two-day case: Suppose that the optimal thing to do on 1-day is to offer a price of p. In the upper left, we see the corresponding allocation curve $a_1(v)$. The bottom left graph shows the revenue curve $R_2(\cdot)$ for 2-day (the thin black curve) and the ironed version $\hat{R}_2(\cdot)$ (the thick blue concave curve). Optimizing for 2-day subject to the inter-day IC constraint $\int_0^v a_1(x)dx \leq \int_0^v a_2(x)dx$ suggests that the most revenue we can get from a deadline $d = 2$ customer is $\hat{R}_2(p)$ on 2-day, which can be done by offering the price of $\underline{p}$ with probability 1/3 and a price of $\overline{p}$ with probability 2/3 (since, in this example, $p = (1/3)\underline{p} + (2/3)\overline{p}$). This yields the pink allocation curve $a_2(v)$ shown in the upper right. The fact that these curves satisfy the inter-day IC constraint follows from the fact that the area of the two grey rectangles shown in the bottom right are equal.

Note that because $\hat{R}_2(\cdot)$ is the least concave upper bound on $R_2(\cdot)$ and by definition of r_2 that $\hat{R}_2(r_2) = R_2(r_2)$. The optimal solution is to set

$$p_1 := \text{argmax}_v R_{12}(v),$$

and then take

$$p_2 := r_2 \text{ if } r_2 \leq p_1 \text{ and } \mathbb{E}(p_2) := p_1 \text{ otherwise},$$

where the randomized case is implemented via the lottery as in the example of Figure 3.2.

The key idea: $R_{12}(v)$ describes the best revenue we can get if we set a price of v for 1-day shipping as shown in Figure 3.3. Since r_2 is the optimal 2-day price, if we are going to set a price above r_2 for 1-day shipping, then the remaining 2-day optimization problem is unconstrained. On the other hand, if the 1-day price is below r_2, then it would constrain the 2-day price via the inter-day IC constraint (3.5), and ironing the 2-day revenue curve may be necessary. This is precisely what the definition of $R_{12}(\cdot)$ in (3.8) does for us. The asymmetry between 1-day and 2-day, specifically the fact that the 1-day curve is never ironed, whereas the 2-day curve is, is a consequence of the inter-day IC constraint (3.5). We generalize this idea in the next section to solve the m-day problem.

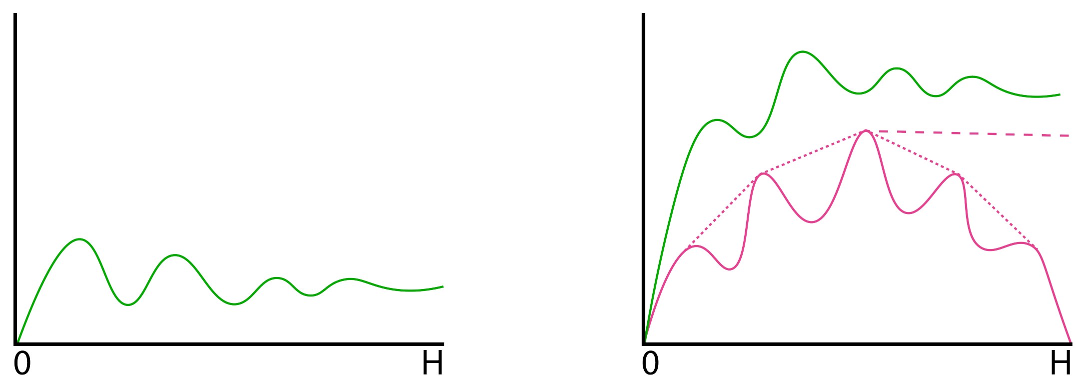


Figure 3.3: Left: The scaled revenue curve for deadline 1, $q_1 R_1(\cdot)$. Right: The pink curve is the scaled revenue curve for deadline 2, $q_2 R_2(\cdot)$. The dotted curve represents the ironed $q_2 \hat{R}_2(\cdot)$. At any possible 1-day price v, the highest pink point at v is the best revenue that can be obtained from 2-day shipping given that a price of v is set for 1-day. This is either the revenue from the ironed curve $q_2 \hat{R}_2(v)$ or, when possible, the revenue of setting a price of r_2, which yields $q_2 R_2(r_2)$. The higher green curve is $R_{12}(\cdot)$, the sum of the green curve from the left and the upper pink envelope, which gives the combined revenue from setting a price of v for 1-day and then doing the best thing for 2-day shipping.

3.4 An optimal allocation rule

3.4.1 Preliminaries

As we discussed regarding the objective, our goal is to choose monotone allocation rules $a_i(v)$ for days $1 \leq i \leq m$ so as to maximize $\sum_{i=1}^{m} q_i \int_0^H \varphi_i(v) f_i(v) a_i(v) dv$.

For a distribution $f_i(\cdot)$ on $[0, H]$ with virtual value function $\varphi_i(\cdot) = v - \frac{1-F_i(v)}{f_i(v)}$, define $\gamma_i(v) := q_i \varphi_i(v) f_i(v)$. Then we aim to choose $a_i(v)$ to maximize $\sum_{i=1}^{m} \int_0^H \gamma_i(v) a_i(v) dv$.

Let $\Gamma_i(v) = \int_0^v \gamma_i(x) dx$. Observe that this function is the negative of the revenue curve, that is, $\Gamma_i(v) = -q_i R_i(v) = -q_i v[1 - F_i(v)]$. [5] Thus, $\Gamma_i(0) = \Gamma_i(H) = 0$ and $\Gamma_i(v) \leq 0$ for $v \in [0, H]$.

Definition 1. For any function Γ, define $\hat{\Gamma}(\cdot)$ to be the lower convex envelope [6] of $\Gamma(\cdot)$. We say that $\hat{\Gamma}(\cdot)$ is *ironed* at v if $\hat{\Gamma}(v) \neq \Gamma(v)$.

Since $\hat{\Gamma}(\cdot)$ is convex, it is continuously differentiable except at countably many points and its derivative is monotone (weakly) increasing.

Definition 2. Let $\hat{\gamma}(\cdot)$ be the derivative of $\hat{\Gamma}(\cdot)$ and let $\gamma(\cdot)$ be the derivative of $\Gamma(\cdot)$.

Claim 1. The following facts are immediate from the definition of lower convex envelope (See Figure 3.4.):

- $\hat{\Gamma}(v) \leq \Gamma(v) \quad \forall v$.
- $\hat{\Gamma}(v_{\min}) = \Gamma(v_{\min})$ where $v_{\min} = \operatorname{argmin}_v \Gamma(v)$. (This implies that there is no ironed interval containing $v_{\min}$.)
- $\hat{\gamma}(v)$ is an increasing function of v and hence its derivative $\hat{\gamma}'(v) \geq 0$ is non-negative for all v.
- If $\hat{\Gamma}(v)$ is ironed in the interval $[\ell, h]$, then $\hat{\gamma}(v)$ is linear and $\hat{\gamma}'(v) = 0$ in (ℓ, h).

We next define the sequence of functions that we will need for the construction:

[5] $\Gamma_i(v) = q_i \int_0^v [x f_i(x) - (1 - F_i(x))] \, dx$. Integrating the first term by parts gives $\int_0^v x f_i(x) \, dx = v F_i(v) - \int_0^v F_i(x) \, dx$. Combining this with the second term yields $\Gamma_i(v) = -q_i v(1 - F_i(v))$.

[6] The *lower convex envelope* of function $f(x)$ is the supremum over convex functions $g(\cdot)$ such that $g(x) \leq f(x)$ for all x. Notice that the lower convex envelope of $\Gamma(\cdot)$ is the negative of the ironed revenue curve $\hat{R}(v)$.

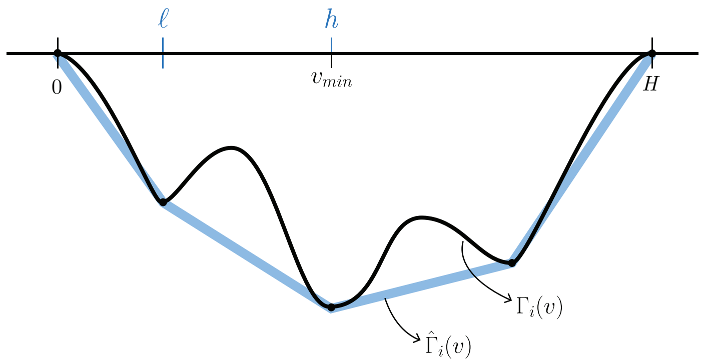


Figure 3.4: The black curve is $\Gamma_i(v)$, and its lower convex envelope $\hat{\Gamma}(v)$ is traced out by the thick light blue line. The curve is ironed in the interval $[\ell, h]$ (among others), so in that interval, $\hat{\Gamma}(v)$ is linear, and thus has second derivative equal to 0.

Definition 3. Let

$$\Gamma_{\geq m}(v) := \Gamma_m(v) \quad \text{and} \quad r_{\geq m} := \text{argmin}_v \Gamma_{\geq m}(v).$$

Inductively, define, for $i := m-1$ down to 1,

$$\Gamma_{\geq i}(v) := \begin{cases} \Gamma_i(v) + \hat{\Gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \Gamma_i(v) + \hat{\Gamma}_{\geq i+1}(r_{\geq i+1}) & v \geq r_{\geq i+1} \end{cases} \quad \text{and} \quad r_{\geq i} := \text{argmin}_v \Gamma_{\geq i}(v).$$

The derivative of $\Gamma_{\geq i}(\cdot)$ is then

$$\gamma_{\geq i}(v) := \begin{cases} \gamma_i(v) + \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \gamma_i(v) & v \geq r_{\geq i+1} \end{cases}.$$

Rewriting this yields

$$\gamma_{\geq i}(v) - \gamma_i(v) = \begin{cases} \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases}. \tag{3.9}$$

Consider when $m = 2$. Since $\Gamma_i(\cdot)$ is the negative revenue curve for i-day shipping scaled by the probability q_i of drawing a customer with deadline i, then when $m = 2$, $\hat{\Gamma}_2(\cdot)$ is the scaled, negative, ironed, revenue curve for i-day shipping, and we aim to minimize it.

Now, we can observe that $\Gamma_{\geq 1}(v)$ is precisely the revenue from setting a price of v for 1-day shipping and optimizing the revenue for 2-day shipping constrained by the price of v set for 1-day shipping. Under this constraint, the best revenue for 2-day shipping is attained by implementing the optimal price from the ironed revenue curve for 2-day shipping that is at most v (using a lottery if needed). Then deadline 2 customers contribute revenue $-1 \cdot \min_{p_2 \leq v} \hat{\Gamma}_2(p_2)$. Observe the pink curve in Figure 3.3: since $\hat{\Gamma}_2(\cdot)$ is concave, the minimum is achieved at $\min\{v, r_2\}$. This is exactly what $\Gamma_{\geq 1}(v)$ accounts for: if $v \leq r_2$, then we get the revenue from setting a price of v for 1-day shipping and the ironed revenue of v for 2-day shipping (possibly via lottery), earning $\Gamma_1(v) + \hat{\Gamma}_2(v)$. If $v \geq r_2$, then we get the revenue from setting a price of v for 1-day shipping and from setting the price of r_2 for 2-day shipping.

For the general case, intuitively, these combined curves account for the fact that when i-day shipping's price is low enough to interfere (with respect to the inter-day IC constraint) with the prices that we would like to set for options $i+1$ through m, we need to consider the problem of setting all of these prices together. However, when i-day's price is high enough not to interfere with the later days, we can just use the optimal choice on days $i+1$ through m (from $\Gamma_{\geq i+1}(r_{\geq i+1})$) and worry about i-day shipping separately. They also take into account the ironing needed to ensure incentive compatibility.

We can draw an analogy to the ironing in Myerson's optimal auction for irregular distributions. Using the ironed curves ensures incentive compatibility and gives an upper bound on the optimal revenue. Myerson shows that this upper bound is in fact achievable using randomization. Similarly, our combined and ironed curves yield upper bounds on the revenue, and we show how to actually achieve these upper bounds by implementing lotteries.

3.4.2 The allocation rule

We define the allocation curves $a_i(\cdot)$ inductively. We use the curve $\Gamma_{\geq i}(\cdot)$ and the constraint from the $(i+1)$-day allocation rule to achieve exactly the revenue that the $\Gamma_{\geq i}(\cdot)$ curves suggest. We will show later that they are optimal. Each allocation curve is piecewise constant. For 1-day shipping, set

$$a_1(v) = \begin{cases} 0 & \text{if } v < r_{\geq 1}, \\ 1 & \text{otherwise.} \end{cases}$$

Suppose that a_{i-1} has been defined for some $i < m$, with jumps at $v_1, \ldots, v_k$, and values $0 = \beta_0 < \beta_1 \leq \beta_2 \ldots \leq \beta_k = 1$. That is,

$$a_{i-1}(v) = \begin{cases} 0 & \text{if } v < v_1, \\ \beta_j & v_j \leq v < v_{j+1} \quad 1 \leq j < k \\ 1 & v_k \leq v. \end{cases}$$

Thus, we can write

$$a_{i-1}(v) = \sum_{j=1}^{k} (\beta_j - \beta_{j-1}) a_{i-1,j}(v)$$

where

$$a_{i-1,j}(v) = \begin{cases} 0 & \text{if } v < v_j \\ 1 & v \geq v_j. \end{cases}$$

Next we define $a_i(v)$.

Definition 4. Let j^* be the largest j such that $v_j \leq r_{\geq i}$. For any $j \leq j^*$, consider two cases:

- $\hat{\Gamma}_{\geq i}(v_j) = \Gamma_{\geq i}(v_j)$, i.e. $\hat{\Gamma}_{\geq i}$ not ironed at v_j: In this case, define

$$a_{i,j}(v) = \begin{cases} 0 & \text{if } v < v_j \\ 1 & otherwise. \end{cases}.$$

- $\hat{\Gamma}_{\geq i}(v_j) \neq \Gamma_{\geq i}(v_j)$: In this case, let
  - $\underline{v}_j :=$ the largest $v < v_j$ such that $\hat{\Gamma}_{\geq i}(v) = \Gamma_{\geq i}(v)$ i.e., not ironed, and
  - $\overline{v}_j :=$ the smallest $v > v_j$ such that $\hat{\Gamma}_{\geq i}(v) = \Gamma_{\geq i}(v)$ i.e., not ironed.

  Let $0 < \delta < 1$ such that

$$v_j = \delta \underline{v}_j + (1 - \delta)\overline{v}_j.$$

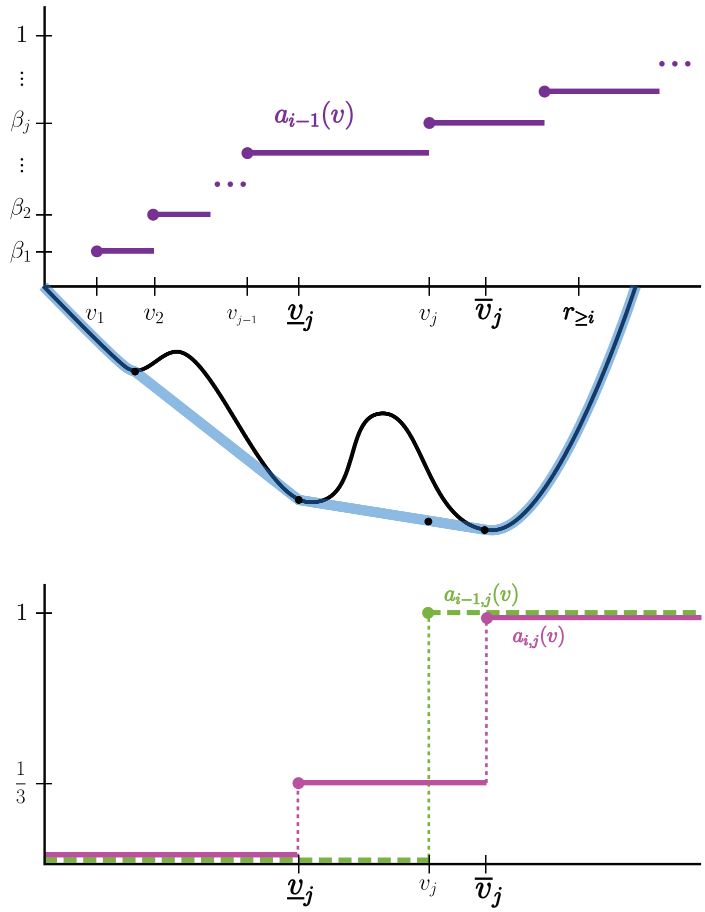


Figure 3.5: This figure shows an example allocation curve $a_{i-1}(v)$ in purple, and illustrates some aspects of Definition 4. The curves $\Gamma_{\geq i}(v)$ and $\hat{\Gamma}_{\geq i}(v)$ are shown directly below the top figure. In this case, $r_{\geq i} \in [v_{j+1}, v_{j+2})$, so $j^* = j + 1$. The bottom figure shows how $a_{i,j}(v)$ is constructed from $a_{i-1,j}(v)$.

Then $\hat{\Gamma}_{\geq i}(\cdot)$ is linear between $\underline{v}_j$ and $\overline{v}_j$:

$$\hat{\Gamma}_{\geq i}(v_j) = \delta \Gamma_{\geq i}(\underline{v}_j) + (1 - \delta)\Gamma_{\geq i}(\overline{v}_j).$$

Define

$$a_{i,j}(v) = \begin{cases} 0 & \text{if } v < \underline{v}_j \\ \delta & \underline{v}_j \leq v < \overline{v}_j \\ 1 & \textit{otherwise.} \end{cases}$$

Finally, set $a_i(v)$ as follows:

$$a_i(v) = \begin{cases} \sum_{j=1}^{j^*}(\beta_j - \beta_{j-1})a_{i,j}(v) & \text{if } v < r_{\geq i}, \\ \\ 1 & v \geq r_{\geq i}. \end{cases} \tag{3.10}$$

Remark: In order to continue the induction and define $a_{i+1}(v)$ we need to rewrite $a_i(v)$ in terms of functions $a_{i,j}(v)$ that take only 0/1 values. This is straightforward.

An Alternate Description: Note that it is equivalent to view $a_{i-1}(\cdot)$ as a randomization over prices where a price of v_j is offered with probability $(\beta_j - \beta_{j-1})$. For each possible price v_j on day $i-1$, we select the optimal choice for i-day shipping using the negative revenue curve $\hat{\Gamma}_{\geq i}(\cdot)$. That is, we determine $p_{i,j} = \operatorname{argmin}_{p \leq v_j} \hat{\Gamma}_{\geq i}(p)$, which is equal to the best constrained price less than v_j to set for i-day shipping to earn revenue for all deadlines i through m, and we implement this price $p_{i,j}$. Note that by convexity, $p_{i,j} = \min\{r_{\geq i}, v_j\}$. In the case that $p_{i,j} = v_j$ and $\hat{\Gamma}_{\geq i}(\cdot)$ is ironed at v_j, this price is implemented by randomizing over $\underline{v}_j$ with probability δ and $\overline{v}_j$ with probability $1-\delta$. We multiply each price (or randomized two prices) for i-day shipping with the probability that v_j was offered for $i-1$-day shipping, giving a randomization over prices for i-day shipping as well, resulting in $a_i(\cdot)$.

In Section 3.3 we mention that for $m = 2$ days, when the reserve prices are increasing, that is, $r_1 < r_2$, and both distributions F_1 and F_2 are regular, the optimal mechanism sets a single price. We now see that the optimal auction sets a price of $r_{\geq 1}$ on 1-day. Since F_2 is regular, the negative revenue curve $\Gamma_2(\cdot)$ is convex, so there is no ironing. Then the best 2-day option is always a price. Furthermore, because F_1 is regular and thus $\Gamma_1(\cdot)$ is convex with its minimum $r_1 < r_2$, the minimum $r_{\geq 1}$ of $\Gamma_{\geq 1}$ will occur between r_1 and r_2. Since $r_{\geq 1} \leq r_2$, then the best price on 2-day will be precisely $r_{\geq 1}$, so there will be a single price.

Note that for each possible price on day $i-1$ we could offer as many as two prices on i-day, hence there is a trivial upper bound of 2^{i-1} options for i-day shipping. Saxena et al. [2018] provides a matching lower bound.

In the special case that each combined curve $\Gamma_{\geq i}(\cdot)$ is convex for each i, then the allocation rule will be a deterministic pricing. This occurs when each scales revenue curve is concave, or when $f_i(v)\varphi_i(v)$ is monotone weakly increasing. This condition is called *declining marginal revenues* (DMR), and we discuss it further in Chapter 4. However, even in this case, determining the allocation rule is more complicated than one might imagine at first glance. For example, for three deadlines, the optimal deterministic mechanism can require 1, 2, or 3 distinct prices, and determining how many prices to use and how to set them is non-trivial using standard revenue curve approaches. However, our definition of the curves $\Gamma_{\geq i}(\cdot)$ makes determining the optimal pricing immediate.

Lemma 1. *The allocation curves $a_i(\cdot)$, for $1 \leq i \leq m$, are monotone increasing from 0 to 1 and satisfy the inter-day IC constraints (3.5). Moreover, each $a_i(\cdot)$ changes value only at points where $\hat{\Gamma}_{\geq i}(\cdot)$ is not ironed.*

Proof. That the allocation curves $a_i(\cdot)$ are weakly increasing, start out at 0, and end at 1 is immediate from the fact that they are convex combinations of the monotone allocation curves $a_{i,j}(\cdot)$. Also, by construction, each $a_i(\cdot)$ changes value only at points where $\hat{\Gamma}_{\geq i}(v)$ is not ironed.

So we have only left to verify that

$$\int_0^v a_{i-1}(x)dx \leq \int_0^v a_i(x)dx.$$

From the discussion above, for $v \leq r_{\geq i}$, we have

$$a_{i-1}(v) = \sum_{j=1}^{j^*}(\beta_j - \beta_{j-1})a_{i-1,j}(v) \quad \text{and} \quad a_i(v) = \sum_{j=1}^{j^*}(\beta_j - \beta_{j-1})a_{i,j}(v)$$

since $a_{i-1,j}(v) = 0$ for $v \leq r_{\geq i}$ and $j > j^*$. Thus, it suffices to show that for each $j \leq j^*$ and $v \leq r_{\geq i}$

$$\int_0^v a_{i-1,j}(x)dx \leq \int_0^v a_{i,j}(x)dx.$$

If $\hat{\Gamma}_{\geq i}$ is not ironed at v_j, then this is an equality. Otherwise, for $v \leq v_j$, the left hand side is 0 and the right hand side is nonnegative. For $v_j \leq v \leq \overline{v}_j$ the left hand side is $(v - v_j)$, whereas the right hand side is $\delta(v - \underline{v}_j)$. Rearranging the inequality $v_j = \delta\underline{v}_j + (1-\delta)\overline{v}_j \geq \delta\underline{v}_j + (1-\delta)v$ implies that $v - v_j \leq \delta(v - \underline{v}_j)$. This completes the proof that (3.5) holds.

Notice that $\int_0^v a_{i-1,j}(x)dx = \int_0^v a_{i,j}(x)dx$ for $v < \underline{v}_j$ and $v > \overline{v}_j$, so $a_{i-1}(v) = a_i(v)$ unless $\Gamma_{\geq i}$ is ironed at v, or $v \geq r_{\geq i}$. We will use this fact in the proof of Claim 4 below. □

3.5 Proof of optimality

In this section, we prove that the allocation rules and pricing of the previous section are optimal. To this end, we formulate our problem as an (infinite) linear program. We discussed the objective and constraints of the primal program in Section 3.2, and we have already shown above that our allocation rules are feasible for the primal program. We then construct a dual program, and a feasible dual solution for which complementary slackness holds. This implies strong duality holds, and thus, that our solution is optimal.

3.5.1 The linear programming formulation

Recall the definitions from Section 3.2: The function $\gamma_i(v)$ is the derivative of $\Gamma_i(v) = \int_0^v q_i\varphi_i(x)f_i(x)\,dx$, where $\varphi_i(v) = v - \frac{1-F_i(v)}{f_i(v)}$ is the i-day virtual value function and q_i is the fraction of bidders with deadline i. Similarly $\hat{\gamma}_i(v)$ is the derivative of $\hat{\Gamma}_i(v)$. We use $[m]$ to denote the set of integers $\{1,\ldots,m\}$.

The Primal

Variables: $a_i(v)$, for all $i \in [m]$, and all $v \in [0,H]$.

$$\text{Maximize } \sum_{i=1}^{m}\int_0^H a_i(v)\gamma_i(v)dv$$

Subject to

$$\int_0^v a_i(x)dx - \int_0^v a_{i+1}(x)dx \le 0 \qquad \forall i \in [m-1] \quad \forall v \in [0,H] \qquad \text{(dual variables } \alpha_i(v)\text{)}$$

$$a_i(v) \le 1 \qquad \forall i \in [m] \quad \forall v \in [0,H] \qquad \text{(dual variables } b_i(v)\text{)}$$

$$-a_i'(v) \le 0 \qquad \forall i \in [m] \quad \forall v \in [0,H] \qquad \text{(dual variables } \lambda_i(v)\text{)}$$

$$a_i(v) \ge 0 \qquad \forall i \in [m] \quad \forall v \in [0,H].$$

Note that $a_i'(v)$ denotes $\frac{d}{dv}a_i(v)$.

The Dual

Variables: $b_i(v), \lambda_i(v)$, for all $i \in [m]$, and all $v \in [0,H]$, $\alpha_i(x)$ for $i \in [m-1]$ and all $x \in [0,H]$.

$$\text{Minimize } \int_0^H [b_1(v) + \cdots + b_m(v)]\, dv$$

Subject to

$$b_1(v) + \lambda_1'(v) + \int_v^H \alpha_1(x)dx \geq \gamma_1(v) \qquad \forall v \in [0,H] \text{ (primal var } a_1(v))$$

$$b_i(v) + \lambda_i'(v) + \int_v^H \alpha_i(x)dx - \int_v^H \alpha_{i-1}(x)dx \geq \gamma_i(v) \qquad \forall v \in [0,H], i = 2,\ldots,m-1$$

$$\text{(primal var } a_i(v))$$

$$b_m(v) + \lambda_m'(v) - \int_v^H \alpha_{m-1}(x)dx \geq \gamma_m(v) \qquad \forall v \in [0,H] \text{ (primal var } a_m(v))$$

$$\lambda_i(H) = 0 \qquad \forall i \in [m]$$

$$\alpha_i(v) \geq 0 \qquad \forall v \in [0,H], i \in [m-1]$$

$$b_i(v), \lambda_i(v) \geq 0 \qquad \forall i \in [m] \forall v \in [0,H].$$

Note that $\lambda_i'(v)$ denotes $\frac{d}{dv}\lambda_i(v)$.

3.5.2 Conditions for strong duality

As long as there are feasible primal and dual solutions satisfying the following conditions, strong duality holds. See Appendix A.1 for a proof that these conditions are sufficient.

$$a_i(v) > 0 \;\Rightarrow\; \lambda_i(v) \text{ continuous at } v \quad i \in [m] \tag{3.11}$$
$$a_i(v) < 1 \;\Rightarrow\; b_i(v) = 0 \quad i \in [m] \tag{3.12}$$
$$a_i'(v) > 0 \;\Rightarrow\; \lambda_i(v) = 0 \quad i \in [m] \tag{3.13}$$
$$\int_0^v a_i(x)dx < \int_0^v a_{i+1}(x)dx \;\Rightarrow\; \alpha_i(v) = 0 \quad i \in [m-1] \tag{3.14}$$
$$b_i(v) + \lambda_i'(v) + \int_v^H \alpha_i(x)dx - \int_v^H \alpha_{i-1}(x)dx > \gamma_i(v) \;\Rightarrow\; a_i(v) = 0 \quad i = 2, \ldots, m-1 \tag{3.15}$$
$$b_1(v) + \lambda_1'(v) + \int_v^H \alpha_1(x)dx > \gamma_1(v) \;\Rightarrow\; a_1(v) = 0 \tag{3.16}$$
$$b_m(v) + \lambda_m'(v) - \int_v^H \alpha_{m-1}(x)dx > \gamma_m(v) \;\Rightarrow\; a_m(v) = 0 \tag{3.17}$$

We allow $a_i'(v) \in \mathbb{R} \cup \{+\infty\}$, otherwise we could not even encode a single-price auction.[7]

3.5.3 The proof

Theorem 2. *The allocation curves presented in Subsection 3.4.2 are optimal, that is, obtain the maximum possible expected revenue.*

Proof. To prove the theorem, we verify that there is a setting of feasible dual variables for which all the conditions for strong duality hold. To this end, set the variables as follows:

$$\lambda_i(v) = \Gamma_{\geq i}(v) - \hat{\Gamma}_{\geq i}(v) \tag{3.18}$$
$$b_i(v) = \begin{cases} 0 & v < r_{\geq i} \\ \hat{\gamma}_{\geq i}(v) & v \geq r_{\geq i} \end{cases} \tag{3.19}$$
$$\alpha_i(v) = \begin{cases} \hat{\gamma}'_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases} \tag{3.20}$$

[7] In particular, $a_i(v)$ may have (countably many) discontinuities, in which points $a_i'(v) = +\infty > 0$. However, in our proof of optimality $a_i'(v)$ appears only as a factor of the product $a_i'(v)\lambda_i(v)$. Every time $a_i'(v) = +\infty$, the corresponding dual value of $\lambda_i(v)$ is 0—by condition (4.3). See also Appendix A.1.

The dual variables are selected precisely to satisfy complementary slackness conditions and therefore ensure optimality.

The dual variable $\lambda_i(\cdot)$ corresponds to the monotonicity constraint on $a_i(\cdot)$ in the primal. Since $\Gamma_{\geq i}(\cdot)$ is the curve used to set a_i, it is intuitive the dual variable $\lambda_i(v)$ corresponds to how much we needed to iron $\Gamma_{\geq i}(v)$ for $a_i(\cdot)$ to be monotone at v.

There are m constraints (other than non-negativity) in the dual program, one corresponding to each deadline. We set them so that the constraint for m-day shipping is satisfied with equality. We can then add constraints m and $m-1$, and the remaining set of $m-1$ constraints correspond precisely to an $m-1$ deadline problem. Herein lies the basis for the induction.

From Claim 1, it follows that $\lambda_i(v), \alpha_i(v) \geq 0$ for all v and i. Since $r_{\geq i}$ is the minimum of $\hat{\Gamma}_{\geq i}(\cdot)$, we have $\hat{\gamma}_{\geq i}(r_{\geq i}) = 0$. Moreover, since $\hat{\gamma}_{\geq i}(\cdot)$ is increasing, $b_i(v) \geq 0$ for all v and i.

Taking the derivative of (3.18), and using Equation (3.9), we obtain:

$$\gamma_i(v) - \lambda_i'(v) = \begin{cases} \hat{\gamma}_{\geq i}(v) - \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \hat{\gamma}_{\geq i}(v) - 0 & v \geq r_{\geq i+1} \end{cases} \tag{3.21}$$

$$\gamma_m(v) - \lambda_m'(v) = \hat{\gamma}_m(v) \tag{3.22}$$

Also, using (3.20) and the fact that $\hat{\gamma}_{\geq i+1}(r_{\geq i+1}) = 0$, we get:

$$A_i(v) := \int_v^H \alpha_i(x)\, dx = \begin{cases} -\hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases} \tag{3.23}$$

Condition (3.11) from Section 3.5.2 holds since $\Gamma_{\geq i}(v)$ and $\hat{\Gamma}_{\geq i}(v)$ are both continuous functions. The proofs of all remaining conditions for strong duality from Section 3.5.2 can be found below. □

Claim 2. Condition (4.2): For all i and v, $a_i(v) < 1 \implies b_i(v) = 0$.

Proof. If $a_i(v) < 1$, then $v < r_{\geq i}$, so by construction, $b_i(v) = 0$. □

Claim 3. Condition (4.3): For all i and v, $a_i'(v) > 0 \implies \lambda_i(v) = 0$.

Proof. From Subsection 3.4.2, $a_i'(v) > 0$ only for unironed values of v, at which $\lambda_i(v) = 0$. □

Claim 4. Condition (4.4): For all i and v, $\int_0^v a_i(x)dx < \int_0^v a_{i+1}(x)dx \implies \alpha_i(v) = 0$.

Proof. As discussed at the end of the proof of Lemma 1, $\int_0^v a_i(x)dx = \int_0^v a_{i+1}(x)dx$ unless $\Gamma_{\geq i+1}$ is ironed at v, or $v \geq r_{\geq i}$. In both of these cases $\alpha_i(v) = 0$ (by part 4 of Claim 1 and Definition 3.20, respectively). □

Claim 5. Conditions (4.5)- (3.17) and dual feasibility: For all i and v, $a_i(v) > 0 \implies$ the corresponding dual constraint is tight, and the dual constraints are always feasible.

Proof. Rearrange the dual constraint $b_i(v) + A_i(v) - A_{i-1}(v) + \lambda_i'(v) \geq \gamma_i(v)$ to

$$b_i(v) - A_{i-1}(v) \geq \gamma_i(v) - \lambda_i'(v) - A_i(v).$$

Fact 1: For $i \in [m-1]$, $\gamma_i(v) - \lambda_i'(v) - A_i(v) = \hat{\gamma}_{\geq i}(v)$ for all v. To see this use (3.21) and (3.23):

$$\gamma_i(v) - \lambda_i'(v) = \begin{cases} \hat{\gamma}_{\geq i}(v) - \hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ \hat{\gamma}_{\geq i}(v) - 0 & v \geq r_{\geq i+1} \end{cases} \qquad A_i(v) = \begin{cases} -\hat{\gamma}_{\geq i+1}(v) & v < r_{\geq i+1} \\ 0 & v \geq r_{\geq i+1} \end{cases}$$

Fact 2: For $i \in \{2, \ldots, m\}$, $b_i(v) - A_{i-1}(v) = \hat{\gamma}_{\geq i}(v)$ for all v.

$$b_i(v) = \begin{cases} 0 & v < r_{\geq i} \\ \hat{\gamma}_{\geq i}(v) & v \geq r_{\geq i} \end{cases} \qquad -A_{i-1}(v) = \begin{cases} \hat{\gamma}_{\geq i}(v) & v < r_{\geq i} \\ 0 & v \geq r_{\geq i} \end{cases}$$

Hence for $i \in \{2, \ldots, m-1\}$, $b_i(v) - A_{i-1}(v) = \gamma_i(v) - \lambda_i'(v) - A_i(v)$ for all v.

For $i = m$, since $\gamma_{\geq m} = \gamma_m$, and $\gamma_m(v) - \lambda_m'(v) = \hat{\gamma}_m(v)$. Combining this with Fact 2 above, we get that $b_m(v) - A_{m-1}(v) + \lambda_m'(v) = \gamma_m(v)$ for all v.

Finally, for $i = 1$, using Fact 1, for $v < r_{\geq 1}$, we get

$$b_1(v) = 0 \geq \hat{\gamma}_{\geq 1}(v) = \gamma_1(v) - \lambda_1'(v) - A_1(v)$$

which is true for $v < r_{\geq 1}$. For $v \geq r_{\geq 1}$, we get

$$b_1(v) = \gamma_{\geq 1}(v) = \gamma_1(v) - \lambda_1'(v) - A_1(v),$$

so the dual constraint is tight when $a_1(v) > 0$ as this starts at $r_{\geq 1}$. □

The above claims prove that this dual solution satisfies feasibility and all complementary slackness and continuity conditions from Section 3.5.2 hold.

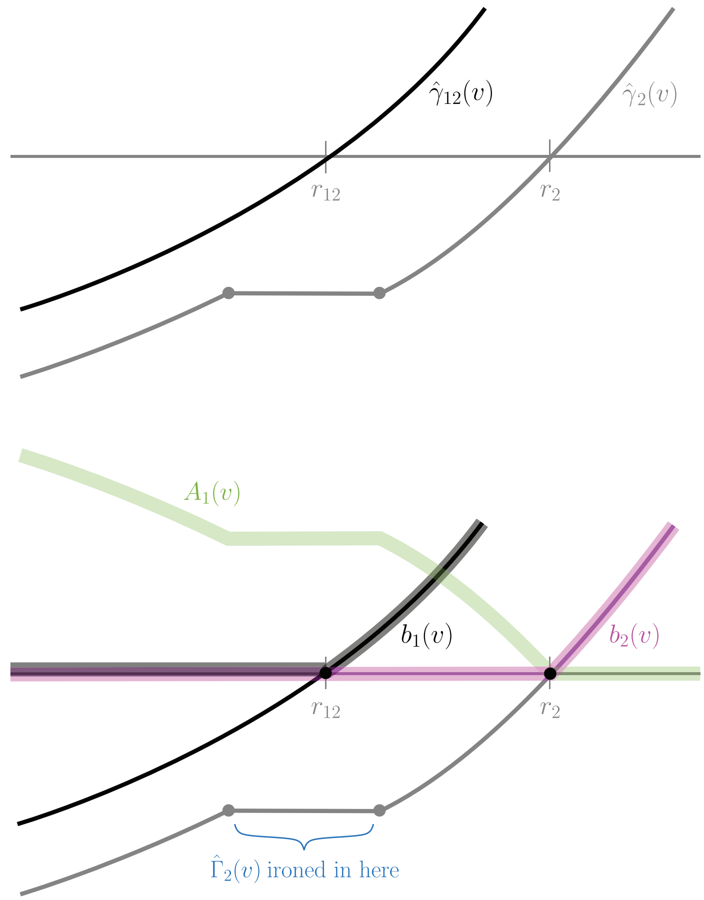


Figure 3.6: This figure illustrates what some of the dual variables might be for the case of two days when $r_{\geq 1} < r_2$. The upper figure plots the functions $\hat{\gamma}_{12}(v)$ and $\hat{\gamma}_2(v)$, and the lower figure shows $b_1(v)$ in dark grey, $b_2(v)$ in pink and $A_1(v) = \int_v^H \alpha_1(x)dx$ in green. Note that up to r_2, the function $A_1(v) = -\hat{\gamma}_2(v)$.

3.6 Closed-Form Virtual Values

By taking the partial Lagrangian of our primal from subsection 3.5.1, we can view the optimal mechanism as an expected virtual welfare maximizer. By plugging in our closed-form dual variables, we produce closed-form virtual value functions. We use our above approach combined with the approximation of Cai et al. [2016], only used for optimal revenue instead of approximation.

We multiply each constraint (aside from feasibility) by its dual variable and move it into the objective function, minimizing over these non-negative dual variables. Recall that basics of Lagrangian duality are outlined in Section 2.4. The resulting partial Lagrangian primal is:

$$\max_{a_i(v)\in[0,1]} \quad \min_{\lambda_i(v),\alpha_i(v)\geq 0} \mathcal{L}(a,\lambda)$$

where

$$\mathcal{L}(a,\lambda) := \sum_{i=1}^{m}\int_0^H a_i(v)\gamma_i(v)dv + \sum_{i=1}^{m-1}\int_0^v \alpha_i(v)\left[\int_0^v a_{i+1}(x)dx - \int_0^v a_i(x)dx\right]dv \\ + \sum_{i=1}^{m}\int_0^v \lambda_i(v)\left[a_i'(v)\right]dv$$

Recall that $A_i(v) = \int_v^H \alpha_i(x)dx$. Then, using this notation, as well as integration by parts on the λ terms, and aggregating the a terms, we can rewrite $\mathcal{L}$ as follows. (This is similar to the steps we take in the proof of strong duality in Appendix A.1.)

$$\begin{aligned}
\mathcal{L}(a,\lambda) &= \int_0^H \left[a_1(v)\gamma_1(v) + \lambda_1(v)a_1'(v) - a_1(v)A_1(v)\right]dv \\
&\quad + \sum_{i=2}^{m-1}\int_0^H \left[a_i(v)\gamma_i(v) + \lambda_i(v)a_i'(v) + a_i(v)A_{i-1}(v) - a_i(v)A_i(v)\right]dv \\
&\quad + \int_0^H \left[a_m(v)\gamma_m(v) + \lambda_m(v)a_m'(v) + a_m(v)A_{m-1}(v)\right]dv \\
&= \int_0^H f_1(v)a_1(v)\left[\frac{\gamma_1'(v)}{f_1(v)} - \frac{\lambda_1'(v)}{f_1(v)} - \frac{A_1(v)}{f_1(v)}\right]dv \\
&\quad + \sum_{i=2}^{m-1}\int_0^H f_i(v)a_i(v)\left[\frac{\gamma_i'(v)}{f_i(v)} - \frac{\lambda_i'(v)}{f_i(v)} - \frac{A_i(v)}{f_i(v)} + \frac{A_{i-1}(v)}{f_i(v)}\right]dv \\
&\quad + \int_0^H f_m(v)a_m(v)\left[\frac{\gamma_m'(v)}{f_m(v)} - \frac{\lambda_m'(v)}{f_m(v)} + \frac{A_{m-1}(v)}{f_m(v)}\right]dv
\end{aligned}$$

This gives that $\mathcal{L}$ is equal to expected virtual welfare for the following virtual values:

$$\phi_i(v) = \frac{\gamma_i'(v)}{f_i(v)} - \frac{\lambda_i'(v)}{f_i(v)} - \frac{A_i(v)}{f_i(v)} + \frac{A_{i-1}(v)}{f_i(v)} \qquad \text{for } i \in \{2, \ldots, m-1\},$$

$$\phi_1(v) = \frac{\gamma_1'(v)}{f_1(v)} - \frac{\lambda_1'(v)}{f_1(v)} - \frac{A_1(v)}{f_1(v)}, \qquad \text{and} \qquad \phi_m(v) = \frac{\gamma_m'(v)}{f_m(v)} - \frac{\lambda_m'(v)}{f_m(v)} + \frac{A_{m-1}(v)}{f_m(v)}.$$

Recall that our partial Lagrangian primal is of the form

$$\max_{a_i(v)\in[0,1]} \quad \min_{\lambda_i(v),\alpha_i(v)\geq 0} \mathcal{L}(a,\lambda) = \max_{a_i(v)\in[0,1]} \quad \min_{\lambda_i(v),\alpha_i(v)\geq 0} \sum_{i=1}^{m} f_i(v)a_i(v)\phi_i(v).$$

Note that ϕ depends on λ, α, and we must find the variables that minimize these functions. We plug in our optimal dual variables from our closed-form solution to FedEx, giving closed-form virtual values:

$$\phi_1 = \hat{\gamma}_1(v)/f_1(v) \qquad \text{and} \qquad \phi_i = \begin{cases} 0 & v < r_{\geq i} \\ \hat{\gamma}_{\geq i}(v)/f_i(v) & v \geq r_{\geq i} \end{cases} \qquad \text{for } i \in \{2, \ldots, m\}.$$

Then, the allocation rule that maximizes expected virtual welfare for these virtual value functions is precisely the optimal mechanism. Note that because our closed-form solution exists, strong duality holds, so the Lagrangian primal is not relaxed.

3.7 Interdimensional Settings

Many other natural problems fall into this category of "interdimensional" as well. Consider a buyer with a value for an item and a private budget b which is the most that he can pay [Devanur and Weinberg, 2017]. Or, suppose a buyer has a value v for each unit of an item up to some private demand capacity d [Devanur et al., 2017].

We highlight now some of the features of the FedEx setting that are common also to the single-minded setting in Chapter 4, as well as these other interdimensional settings. These properties also help to explain why duality techniques gain traction.

- Every allocation rule (which lists, for each (value, deadline) pair, a probability of receiving each of the three items) can be "collapsed" to simply list, for each (value, deadline) pair, a single probability (of receiving a satisfying item).

- Local Incentive Compatibility (IC) constraints imply global IC constraints. That is, any auction satisfying all local IC constraints is also globally IC.
- A payment identity applies: a simple closed form determines payments as a function of the allocation rule.

Of course, these three properties are intertwined: without a collapsible allocation rule, no closed-form payment identity is possible.

There are other commonalities as well. In each of these settings, it has been shown (in this chapter, the subsequent chapter, [DW '17], and [DHP '17]) that the optimal mechanism is deterministic when the marginal distributions satisfy declining marginal revenues (DMR). These works also show that the degree of randomization (the menu complexity) of the optimal mechanism is larger than that in single-dimensional settings, but smaller than in multi-dimensional settings, for the FedEx, budgets, and the single-minded settings.

4 Single-Minded Agents

4.1 Introduction

Consider the problem of selling multiple items to a unit-demand buyer. The fundamental problem underlying much of mechanism design asks how the seller should maximize their revenue. If the items are identical, then the setting is considered *single-dimensional.* In this case, seminal work of Myerson [1981] completely resolves this question with an exact characterization of the optimal mechanism. The optimal mechanism is a simple take-it-or-leave-it price, and the fact that there are multiple items versus just one is irrelevant. In contrast, if the items are heterogenous, then the setting is *multi-dimensional* and, unlike the single-dimensional setting, optimal mechanisms are no longer tractable in any sense: numerous recent works identify various undesirable properties [Manelli and Vincent, 2007; Briest et al., 2015; Hart and Nisan, 2013; Hart and Reny, 2012; Daskalakis et al., 2013; Daskalakis, Deckelbaum, and Tzamos, 2015].

Chapter 3 identifies an interesting middle-ground. Imagine that the items are neither identical nor heterogeneous, but are instead varying qualities of the same item. To have an example in mind, imagine that you're shipping a package and the items are one-day, two-day, or three-day shipping. You obtain some value v for having your package shipped, but only if it arrives by your deadline (which is one, two, or three days from now). Viewed in the context of a unit-demand buyer, this means that the buyer will always have value v or 0 for every item, and the set of items which yield non-zero value is either $\{1\}$, $\{1, 2\}$, or $\{1, 2, 3\}$ (so we can think of the input as being a two-dimensional distribution over (value, deadline) pairs).

The FedEx Problem is a special case of single-minded valuations: a buyer has a valuation v for a specific subset of items S, and obtains value v if he gets any superset of S, and 0 otherwise. To have an example in mind, imagine that a company offers internet, phone service, and cable TV. You have a value, v, and are interested in either exclusively internet service, internet/phone service, or internet/cable, and so on. If you receive something at least as good as your interest, then you get value v, otherwise you get a value of zero (so we again think of the input distribution as a

This chapter is based on ongoing joint work with Nikhil Devanur, Raghuvansh Saxena, Ariel Schvartzman, and Matt Weinberg [DGSSW '19].

two-dimensional distribution over (value, interest) pairs).

An alternative perspective to single-minded valuations is that there is a partial order on the set of possible interests a buyer may have. The partial order is just the one induced by set inclusion. The FedEx problem has *totally-ordered* items: one-day shipping is at least as good as two-day shipping is at least as good as three-day shipping, and every buyer agrees. In fact, any partial order can be induced from set inclusion, so the two settings are equivalent (see Observation 13 in Section 4.10). It turns out that the partial order view is more useful from a mechanism design perspective, therefore we will use that view for the rest of this chapter.

The following problem can also be interpreted as a partially-ordered setting: Suppose that each buyer has a publicly visible attribute which the seller can use to price discriminate. E.g., the buyer could be a student, a senior, or general-admission. Or, the buyer could be a "prime member" or a "non-prime member." However, buyers with certain attributes can disguise themselves as having other attributes, given by a partial order. For example, a prime member could disguise as a non-prime member, but not vice-versa. Then if item i is a movie ticket redeemable by anyone who can disguise themselves as having attribute i, the items are partially-ordered.

4.1.1 Main Results

Of particular relevance to this chapter is the notion of *menu complexity*: the number of non-trivial options presented to the buyer. Through the lens of menu complexity, Myerson's seminal work shows that the optimal mechanism for single-dimensional settings has menu complexity 1, while Manelli and Vincent [2007]; Daskalakis et al. [2015] show that the optimal mechanism for the multi-dimensional setting might have *uncountable* menu complexity—this holds even for just two items, and even when the item values are drawn independently from absolutely bounded distributions. This dichotomy serves as one fundamental difference between single-dimensional and multi-dimensional settings.

The main results of this chapter are upper and lower bounds showing that for the smallest non-trivial instance of just three partially-ordered items, the menu-complexity of the revenue-optimal mechanism is *unbounded but finite*.[1] This uncovers a new region on the menu complexity spectrum, sharply contrasting with the

[1]This work started as an attempt to generalize the FedEx results to the single minded setting. This chapter is an explanation for *why we failed*.

single-dimensional setting (menu complexity 1), the totally-ordered setting (menu complexity 7, for three items), and the heterogeneous items setting (uncountable).

An *unbounded* lower bound (Theorem 3) implies that as soon as the partial order is non-trivial (i.e., not totally-ordered), optimal mechanisms get considerably richer. In particular, our lower bound implies that optimal mechanisms for single-minded valuations with at least three items have unbounded menu complexity. On the other hand, we show that the menu complexity for the three item case is *always finite* (Theorem 4). Together, this means that for all M, there exists a distribution over (value, interest) pairs such that every optimal mechanism has menu complexity $\geq M$, but for any distribution, there exists an optimal menu of finite menu complexity.

Note that while it is a subject of debate what the true notion of complexity or simplicity should be for mechanisms, our lower bounds are in a setting where there are exactly 3 deterministic outcomes. Thus the source of high menu complexity is randomization, and therefore our results can also be thought of as capturing the level of randomization required by the optimal mechanism. It should be of clear interest that many different natural settings can span such a wide gap of randomization needed in the optimal mechanism.

The main technical takeaway from our results is a thorough understanding of optimal mechanisms for single-minded bidders through broadly applicable tools. Our theorem statements use the language of menu complexity, but only to distinguish among mechanisms with bounded, unbounded, or infinite menu complexity. The main conceptual takeaway is that optimal auctions for single-minded valuations lie in a space of their own: significantly more complex than optimal single-dimensional auctions, or even optimal auctions for totally-ordered valuations, yet significantly more structured than optimal multi-dimensional auctions.

To contrast with the complexities that result when no assumptions are placed on the prior distributions, we also consider the special case where each marginal distribution (of values conditioned on interest) satisfies Declining Marginal Revenues.[2] Here, we provide an explicit description of the optimal mechanism, which is *deterministic* (Theorem 12). That is, each item i (in the language of partial orders, or bundle in the language of single-minded buyers) is offered at some price p_i and

[2]A one-dimensional distribution F satisfies Declining Marginal Revenues (DMR) if $v(1 - F(v))$ is concave. See Devanur et al. [2017] for examples and more discussion. For example, uniform distributions are DMR, along with any distribution of bounded support and monotone non-decreasing density.

Known Menu Complexity Results for Optimal Mechanisms with One Buyer

	One Item	FedEx	Single-Minded, 3 Items	Multi-Unit	Additive
Det. under DMR	N/A	✓	**✓**	✓	N/A
Lower Bound	1	$2^m - 1$	**unbounded**	**unbounded**	uncountable
Upper Bound	1	$2^m - 1$	**finite**	—	uncountable

Bold results are from this work.

the buyer picks whichever item maximizes $v_i - p_i$ (so the menu complexity is at most m for m items). The fact that optimal mechanisms are deterministic subject to DMR matches prior work for totally-ordered settings [Che and Gale 2000; FGKK '16; DW '17; DHP '17].

4.1.2 Additional results

We postpone all details about our proofs to the technical sections, but highlight one result of independent interest that we develop en route. Our problem can be phrased as a continuous linear program, and all of our proofs require reasoning about the dual. In particular, developing our lower bound construction (instances with unbounded menu complexity) consists of two parts: First, we construct a candidate dual λ for which a primal exists satisfying complementary slackness, and for which *every* primal satisfying complementary slackness has menu complexity $\geq M$. Second, we prove that there exists a distribution for which λ is a feasible dual (and combining these two claims means that every optimal mechanism for this input has menu complexity $\geq M$). Analyzing λ through complementary slackness is technically interesting, and captures all of the insight one would hope to gain from the construction. Reverse engineering an instance for which λ is feasible, however, is technically challenging yet unilluminating. On this front, we prove a "Master Theorem," stating essentially that every candidate dual is feasible for some input distribution (Theorem 14). This allows the user (of the theorem) to reason exclusively about primals and duals, letting the Master Theorem map the candidate pair back to an instance for which they are feasible. In some sense, the Master Theorem formally separates the insightful analysis from the tedious parts.

Of course, one should not expect this theorem to hold in general multi-dimensional settings (in particular, one key property that enables our Master Theorem is a "payment identity," which general multi-dimensional settings notoriously lack—this is a further example of how our setting lies in-between single- and multi-dimensional), but the Master Theorem is quite generally applicable for problems in this intermediate range. In addition, because the Master Theorem takes care of guaranteeing that distributions corresponding to some dual will exist, this result also emphasizes the strength of reasoning about duals in similar settings.

Finally, beyond our main results, we prove two additional results using the same tools. First, we apply our lower bound techniques to show that the menu complexity of the Multi-Unit Pricing problem [DHP '17] is also unbounded (Theorem 18 in Section 4.12). Multi-Unit Pricing is also a totally-ordered setting, where the items correspond to copies of a good (item one is one copy, item two is two copies, item three is three copies). The difference from FedEx is that if the buyer is interested in two copies but gets one, they get half their value (versus zero).

4.1.3 Roadmap

We show that the menu complexity of optimal mechanisms for single-minded bidders is unbounded (Theorem 3), but always finite (Theorem 4), even with just $m = 3$ items. We further show that optimal mechanisms are deterministic whenever all marginals satisfy DMR, and provide an explicit construction of the optimal mechanism (Theorem 12). Our "Master Theorem" (Theorem 14) is of independent interest for future work on mechanism design with totally- or partially-ordered items. We also display the applicability of our techniques for the related settings of Multi-Unit Pricing (Theorem 18).

Immediately below, we overview the most related works. In Section 4.3, we provide the minimal preliminaries to get the main ideas behind our results (full preliminaries appear in Section 4.6). In Section 4.4, we highlight the main features of optimal mechanisms for single-minded bidders without yet getting into duality. In Section 4.5 we overview the key duality aspects for our main results. Sections 4.7 and 4.8 prove our main result in the general case. Section 4.9 proves our stronger characterization for the special case of DMR marginals.

4.2 Related Work

The most related line of works has already mostly been discussed. The FedEx Problem considers totally-ordered items (in our language), as does Multi-Unit Pricing and Budgets [Che and Gale, 2000; Fiat et al., 2016; Devanur et al., 2017; Devanur and Weinberg, 2017]. The present chapter is the first to consider partially-ordered items. In terms of techniques, we indeed draw on tools from prior work. All three prior works employ some form of duality. Our approach is most similar to that of Devanur and Weinberg [2017] in that (1) both are the only works to use the analysis from [CDW '16] to characterize optimal mechanisms rather than obtain approximations, and (2) we also perform "dual operations" rather than search for a closed form. However, as the single-minded setting is much more complicated, we extend the techniques to handle this setting.

Also related is a long line of work which aims to characterize optimal mechanisms beyond single-dimensional settings. Owing to the inherent complexity of mechanism design for heterogeneous items, results on this front necessarily consider restricted settings [Laffont et al., 1987; Giannakopoulos and Koutsoupias, 2014; McAfee and McMillan, 1988; Daskalakis et al., 2013, 2015; Haghpanah and Hartline, 2015; Malakhov and Vohra, 2009]. From this set, the most related are Haghpanah and Hartline [2015]; Malakhov and Vohra [2009], who also considered settings where all consumers prefer (e.g.) item a to item b, but there are no substantial technical connections.

There is also a quickly growing body of work regarding the menu complexity of multi-item auctions. Much of this work focuses on settings with heterogeneous buyers [Briest et al., 2015; Hart and Nisan, 2013; Babaioff, Gonczarowski, and Nisan, 2017; Wang and Tang, 2014; Daskalakis et al., 2015; Gonczarowski, 2018]. Very recent work of [SSW '18] considers the menu complexity of approximately optimal mechanisms for the FedEx Problem (for which [FGKK '16] already characterized the menu complexity of exactly optimal mechanisms). On this front, our work places partially-ordered items (where the menu complexity is finite but unbounded) distinctly between totally-ordered items (where the menu complexity is bounded) [FGKK '16], and heterogeneous items (uncountable) [DDT '15]. Previously, no settings with this property were known.

4.3 Preliminaries

In the interest of presentation, we'll provide the minimum preliminaries here for the reader to understand the key ideas. In Section 4.6, we provide full preliminaries, including additional intuition, and covering prior work (such as [FGKK '16; DW '17]). Many of the facts we will use are stated here without proof (proofs are given in Section 4.6).

4.3.1 A Minimal Instance

We focus on the three-item case with items $\mathcal{G} = \{A, B, C\}$ where $A \succ C$ and $B \succ C$, but $A \not\succ B$ and $B \not\succ A$. That is, if a buyer is interested in item C, they are content with A or B. If they are interested in A, they are content only with A (ditto for B). There is a single buyer with a (value, interest) pair (v, G), who receives value v if they are awarded an item $\succeq G$ (that is, $G' \succ G$ or $G' = G$). This is the minimal non-trivial example of a partially-ordered setting. A menu-complexity lower bound for this example applies to any partially-ordered setting that contains an item G with at least two incomparable items that dominate G (which includes every single-minded valuation setting with at least 3 items). For any partially-ordered instance that does *not* contain this structure, i.e., if for every item G, there is at most one other item that immediately dominates G, then the FedEx closed-form solution applies[3], so the menu-complexity of the optimal mechanism is exponential/bounded [FGKK '16; SSW '18].

An instance of the problem consists of a joint probability distribution over $[0, H] \times \mathcal{G}$, where H is the maximum possible value of any bidder for any item.[4] We will use f to denote the density of this joint distribution, with $f_G(v)$ denoting the density at (v, G). We will also use $F_G(v)$ to denote $\int_0^v f_G(w)dw$, and q_G to denote the probability that the bidder's interest is G.

We'll consider (w.l.o.g.) direct truthful mechanisms, where the bidder reports a (value, interest) pair and is awarded a (possibly randomized) item. Further, as observed in Fiat et al. [2016], it is without loss of generality to only consider mechanisms that award bidders their declared item of interest with probability in $[0, 1]$, and all other items with probability 0.[5] For a direct mechanism, we'll define $a_G(v)$ to be

[3] This is not an immediate corollary of their theorem statements, but readers familiar with Fiat et al. [2016] will observe that this follows.

[4] Note that the multi-dimensional instances with uncountable menu complexity are also supported on a compact set: $[0, H]^2$. So our results are not merely a product of compactness.

[5] To see this, observe that the bidder is just as happy to get nothing instead of an item that doesn't dominate their interest. See also that they are just as happy to get their interest item instead of any item that dominates it. It will also make this option no more attractive to any bidder considering

the probability that item G is awarded to a bidder who reports (v, G). Our goal is to find the revenue-optimal allocation rule—$a_G(v)$ defined for all $G \in \mathcal{G}, v \in [0, H]$ with payment determined by the allocation rule—such that the mechanism is incentive-compatible. The menu complexity of a direct mechanism refers to the number of distinct pairs (G, q) such that there exists a v with $a_G(v) = q$.

4.3.2 Incentive Compatibility, Revenue Curves, and Ironing

As observed in [FGKK '16], it is without loss of generality to only consider mechanisms that award bidders their declared item of interest with probability in $[0, 1]$, and all other items with probability 0. Also observed in [FGKK '16] is that Myerson's payment identity holds in this setting as well, and any truthful mechanism must satisfy $p_G(v) = va_G(v) - \int_0^v a_G(w)dw$ (this also implies that the bidder's utility when truthfully reporting (v, G) is $u_G(v) = \int_0^v a_G(w)dw$). This allows us to drop the payment variables, and follow Myerson's analysis. Fiat et al. observe that many of the truthfulness constraints are redundant, and in fact it suffices to only make sure that when the bidder has (value, interest) pair (v, G) they:

- Prefer to tell the truth rather than report any other (v', G). This is accomplished by constraining $a_G(\cdot)$ to be monotone non-decreasing (exactly as in the single-item setting).

- Prefer to tell the truth rather than report any other $(v, G' \in N^+(G))$. By $N^+(G)$, we mean all items G' such that $G' \succeq G$, but there does not exist a G'' with $G' \succeq G'' \succeq G$. This is accomplished by constraining $\int_0^v a_G(w)dw \geq \int_0^v a_{G'}(w)dw$ (as the LHS denotes the utility of the buyer for reporting (v, G) and the RHS denotes the utility of the buyer for reporting (v, G')). Note that this is equivalent to saying that the area under G's allocation curve should be at least as large *at every* v as the area under G''s allocation curve.

All of these constraints together imply that (v, G) also does not prefer to report any other (v', G').[6] We conclude this section with some standard definitions and

misreporting. So starting from a truthful mechanism, modifying it to only award the item of declared interest or nothing cannot possibly violate truthfulness. Note also that this modification maintains optimality, but could impact the menu complexity up to a factor of # items. As we only consider distinctions between bounded, unbounded, and infinite, this is still w.l.o.g.

[6]For example, if (v, G) prefers truthful reporting to reporting (v, G') where $G' \succ G$, and (v, G') prefers truthful reporting to reporting (v', G'), then since (v, G) gets the same utility for reporting

observations.

Definition 5 (Revenue Curve)**.** The *revenue curve* for an item G with CDF F_G is a function R_G that maps a value v to the revenue obtained by posting a price of v, for a single item, when buyer values are drawn from the distribution F_G. Formally, $R_G(v) := v \cdot [1 - F_G(v)]$. We say that a revenue curve is feasible if there exists a distribution that induces it. The *monopoly reserve price* r_G of the revenue curve is $r_G \in \text{argmax}_p R_G(p)$.

Definition 6 (Virtual Value)**.** Myerson's virtual valuation function $\varphi_G(\cdot)$ is defined so that $\varphi_G(v) := v - \frac{1-F_G(v)}{f_G(v)}$. Observe that $R'_G(v) = 1 - F_G(v) - v f_G(v) = -\varphi_G(v) f_G(v)$. When clear from context we will omit the subindex G.

Definition 7 (DMR)**.** We say that a marginal distribution of values F_G satisfies *declining marginal revenues* (DMR) if $R_G(v)$ is concave, or equivalently, if $\varphi_G(v) f_G(v)$ is monotone non-decreasing.

When the marginal distributions do not all satisfy the DMR assumption, we instead need to iron the distribution, an analogue to Myersonian ironing.

Definition 8 (Ironing)**.** The *ironed revenue curve* denoted $\hat{R}(\cdot)$ for a revenue curve $R(\cdot)$ is the least concave upper bound on the revenue curve $R(\cdot)$.[7] A point v is *ironed* if $\hat{R}(v) \neq R(v)$. We say that $[a, b]$ is an *ironed interval* if $\hat{R}(a) = R(a)$, $\hat{R}(b) = R(b)$, and $\hat{R}(v) \neq R(v)$ for all $v \in (a, b)$, where if $v \in (a, b)$, then a and b are the lower and upper endpoint of the ironed interval, respectively.

An ironed revenue curve is depicted in Figure 4.1. By the definition of concavity, if z is ironed, then $\hat{R}(z) = \beta R(a) + (1 - \beta) R(b)$ where $z \in (a, b)$, $\beta a + (1 - \beta) b = z$, and a, b are unironed. Importantly, observe that setting price z to a consumer drawn from F_G yields revenue $R(z) < \hat{R}(z)$. Yet, if we set price a with probability β and b with probability $(1 - \beta)$, we will get revenue $\beta R(a) + (1 - \beta) R(b) = \hat{R}(z)$. Once can

(v, G') as type (v, G') does for truthfully reporting, (v, G) prefers truthful reporting to reporting (v', G').

[7] We emphasize that this work irons the revenue curve with *values* on the x-axis. Classical one-dimensional ironing (to yield Myersonian ironed virtual values) is done on the revenue curve with quantiles on the x-axis.

check that this is precisely the allocation and payment

$$a(v) = \begin{cases} 0 & v < a \\ \beta & v \in [a, b) \\ 1 & v \geq b \end{cases} \quad \text{and} \quad p(v) = \begin{cases} 0 & v < a \\ \beta a & v \in [a, b) \\ \beta a + (1-\beta) b & v \geq b \end{cases}.$$

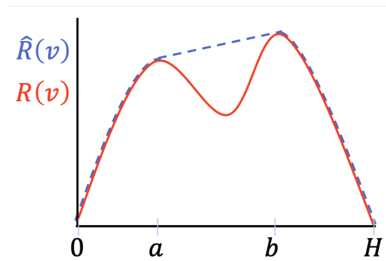


Figure 4.1: For some implicit distribution F, the revenue curve $R(v) = v \cdot [1 - F(v)]$ is depicted, as is the ironed revenue curve, or the revenue curve's least concave upper bound.

4.4 Three Illustrative Examples

In this section, we use three example instances to understand how the optimal mechanisms become increasingly complex, blowing up from deterministic prices to unbounded randomization. We begin with some intuition before diving into examples.

Intuition: Why is single-minded more complex? Consider first a one-item setting that only sells 2-day shipping. Myerson's seminal work proves that the optimal way to sell 2-day shipping in isolation is to post the monopoly reserve price for it. Consider next retroactively adding 1-day shipping into the mix, perhaps because some customers demand 1-day shipping and aren't satisfied with 2-day shipping. Perhaps the distribution of customers demanding 1-day shipping has a higher Myerson reserve than the initial 2-day shipping distribution, in which case it is consistent to set both optimal reserves. Note, however, that a customer who wants their package within 2 days would be content with 1-day shipping. So if instead the 1-day shipping distribution has a *lower* Myerson reserve than 2-day shipping, posting the pair of Myerson reserves is no longer incentive compatible. This complexity arises in the FedEx problem Fiat et al. [2016], and requires considering the constraints imposed on 2-day shipping by 1-day shipping (or vice versa).

Now consider the simplest single-minded valuation setting. The internet service provider (ISP) sells three options: wifi, wifi/cable, and wifi/phone, where wifi/-cable and wifi/phone dominate wifi but are incomparable with each other. If it happens to be that the distribution of consumers who are interested in wifi/cable or wifi/phone both have a higher Myerson reserve than the distribution of consumers who are interested in only wifi,[8] then again the seller can simply offer all three options at their Myerson reserve. However, if this is not the case, further optimization must be done. Importantly, in contrast to the FedEx setting, there's a circular dependency involving these three options which doesn't arise in the totally-ordered case (see examples for further detail). In this way, the IC constraints that govern the mechanism are much more complex in the single-minded setting than in the FedEx setting, and are the reason both for developing much richer techniques and for the much higher degree of randomization that is seen in our results.

Now, we explain what the optimal mechanism looks like for (1) the minimal partially-ordered (single-minded) instance under DMR, (2) the minimal totally-ordered (FedEx) instance without DMR, and (3) the minimal partially-ordered instance without DMR.

Three Partially-Ordered Items under DMR. We begin with the special case where the marginal distributions for each item satisfy DMR. Recall that this implies that the marginal revenue curves for each item are concave, and thus do not require ironing. We show how to derive the optimal item pricing (but a proof that this is indeed optimal is deferred to Section 4.9 as part of the general DMR case). Our instance is again that where C is the worst item (e.g. wifi) and A and B are incomparable (e.g. wifi/cable and wifi/phone).

Let's start by considering what price we would set for item A if we had already set price p_C for item C. (Note that whatever price we set for item B has no effect, as A and B are incomparable.) Observe that our revenue from setting any price p_A is just $p_A \cdot [1 - F_A(p_A)]$, so ideally we would just set price $r_A := \arg\max_p \{p \cdot [1 - F_A(p)]\}$.

[8] Recall the example from Chapter 3 in which a one-dimensional distribution D stochastically dominates D' yet has a lower reserve: D' is uniform over the set $\{1, 10\}$ and has reserve 10; D is uniform over the set $\{9, 10\}$ with reserve 9.

If $r_A \geq p_C$, this doesn't violate any IC constraints. Indeed, consumers with interest C will prefer to pay $p_C \leq r_A$ to get item C rather than item A. If $r_A < p_C$, however, setting price r_A will violate IC, as now consumers with interest C would strictly prefer to report interest in item A instead. This constrains us to set a price for A that is at least p_C. Observe that, because $R_A(\cdot)$ is concave, the revenue-maximizing price to set that is at least p_C (which is $> r_A$) is $p_A := p_C$. Hence, we can define the revenue curve $\bar{R}_A(\cdot)$ to describe the revenue we can get from selling item A as a function of p_C:

$$\bar{R}_A(p_C) = \begin{cases} R_A(r_A) & p_C \leq r_A \\ R_A(p_C) & p_C > r_A \end{cases}.$$

The same definition holds for $\bar{R}_B(\cdot)$. Now, we can find the price to set for item C that optimizes the impact on all three items by simply finding the p maximizing $R_{ABC}(p) := R_C(p) + \bar{R}_A(p) + \bar{R}_B(p)$ (depicted in Figure 4.2). Picking p_C as such, and then setting $p_A := \max\{r_A, p_C\}$, $p_B := \max\{r_B, p_C\}$ is the optimal pricing. The (challenging) remaining step is to prove that in fact this is optimal even among randomized mechanisms. The duality theory previously hinted at is key in this step, but we postpone these details for now. Importantly, note that this claim requires the DMR assumption (so proving it will certainly be technically involved)—without it, there might be a better randomized mechanism.

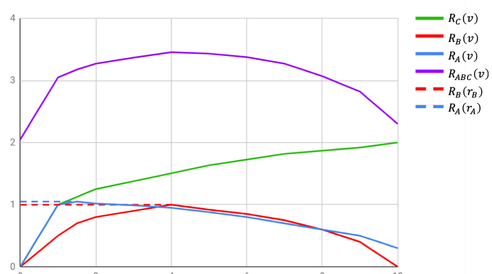


Figure 4.2: The construction of R_{ABC} with $R_A, \bar{R}_A, R_B, \bar{R}_B$, and R_C illustrated as well.

Two Items without DMR (FedEx). In this example, there are only two items, A and C with $A \succ C$. In this case, we'll think about first setting the price for A, and understanding how it constrains our choices for C. If we set price p_A for item A, then we are constrained to give every type (v, C) interested in item C utility at least $v - p_A$. Again, if $r_C \leq p_A$, we should just set price r_C on item C. However, if $r_C > p_A$, *without the DMR assumption*, it's unclear what the best price to set should be. Indeed, it could be that some price $p_C \ll p_A$ generates more revenue than p_A as $R_C(\cdot)$ is not necessarily concave. Note, however, that the ironed revenue curve $\hat{R}_C(\cdot)$ *is* concave. So $\arg\max_{p_C \leq p_A}\{\hat{R}_C(p_C)\} = \min\{r_C, p_A\}$. It's unclear exactly what to make of this, but one hope (that turns out to be correct), is that the optimal scheme for item C, conditioned on p_A, is to set *expected price* $p_C := \min\{r_C, p_A\}$ via the allocation rule defined as in Definition 8. It is not obvious that such an allocation rule satisfies IC, but straight-forward calculations confirm that indeed it does. Similarly to the previous example, we can now define:

$$\bar{R}_C(p_A) = \begin{cases} \hat{R}_C(p_A) & p_A < r_C \\ R_C(r_C) & p_A \geq r_C \end{cases} \qquad \text{and} \qquad R_{AC}(p_A) = R_A(p_A) + \bar{R}_C(p_A).$$

This construction is depicted in Figure 4.3. Figure 4.4 gives some intuition as to why it is indeed incentive compatible to set the proposed allocation rule for item C (but the goal of this section is not to provide complete proofs). It is now clear that, among all options which set a deterministic price for item A, and implement an expected price on the ironed revenue curve for item C, the above procedure is optimal. What is not clear is why this procedure is optimal over all possible menus for item C, or even why a randomized menu for item A can't perform better. Indeed, the same duality theory referenced previously takes care of this.

This example perhaps also gives intuition for the menu complexity upper bound of $2^m - 1$ for FedEx. Repeating this process for another totally-ordered item, each option offered to buyers with interest C could be "split" into at most two new options to be offered to buyers with interest $D \prec C$.

Three Partially-Ordered Items without DMR. In our first example, we reasoned about how our decision for item C constrains which prices to set for items A and B. In our second example, we reasoned about how our decision for item A constrains prices to set for item C. We presented the opposite direction (1) to present both types of arguments and (2) because this direction is necessary without the DMR assumption. For partially-ordered items, however, we really can only reason about how decisions for item C constrain prices for A and B. The reason is that in order to know how p_A constrains our options for item C, we also need to know p_B. Indeed, only $\min\{p_A, p_B\}$ matters for constraining C. So we would need to know p_B to

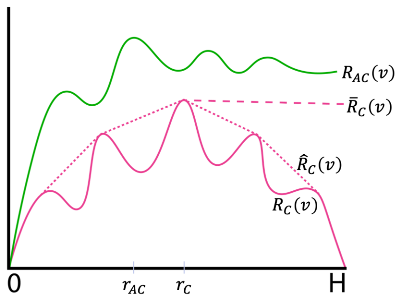


Figure 4.3: A worse-to-better item revenue curve for the FedEx setting that determines the optimal mechanism even without DMR.

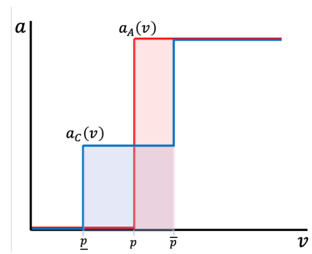


Figure 4.4: Utility for items A and C are equal for $v \leq \underline{p}$ and $v \geq \underline{p}$, but for $v \in (\underline{p}, \overline{p})$, the randomized option provides more utility.

know whether a proposed p_A is imposing a new constraint or not. This results in an impasse for this approach: this partial order requires us to reason about C's price first, but without DMR, we must reason about A and B first. However, this is only intuition as to why this setting becomes more complicated. In Subsection 4.5.2, we explain why it is that the IC constraints can cause the randomization to get so unwieldy, and Subsection 4.5.4 cements this with an example.

Note, however, that we can still reason as we previously did about the optimal item pricing. If, as in the first example, we define $\bar{R}_A(p_C)$ to be the revenue from selling item A at the optimal price that exceeds p_C, and $\bar{R}_B(p_C)$ similarly for item B, then $R_{ABC}(p_C) := R_C(p_C) + \bar{R}_B(p_C) + \bar{R}_A(p_C)$ accurately defines the revenue we get from all three items by setting price p_C on item C, and setting the optimal prices for A and B conditioned on this.

Take the following example marginal distributions and corresponding marginal revenue curves, depicted in Figures 4.5 and 4.6, which do not satisfy DMR.

The optimal deterministic price to set is 8 on item C, which will result in prices of 8 on items A and B as well. This gives $R_{ABC} = 3.155$. However, there is a better mechanism this time, with a good deal of randomization. Working out the exact numbers is quite messy, which further motivates the duality approach we take. But the idea is that the interaction between the three items can get quite unwieldy: improving the option offered for C relaxes constraints on A, which may increase the utility received by purchasing item A. This can in turn loosen the constraints on item B (because now item A is preferable to item B), which may in turn cause

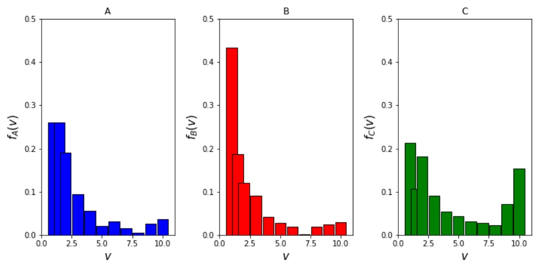


Figure 4.5: The probability densitives for items A, B, and C, that do not satisfy DMR.

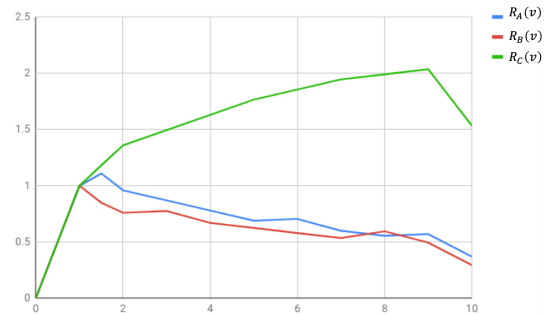


Figure 4.6: The corresponding non-DMR revenue curves.

higher-revenue options for item C to be viable, etc. Coming up with an explicit solution to cope with this circular reasoning is fairly intractable, but fortunately it is relatively tractable to describe potential optimal solutions through Lagrangian duality and complementary slackness. The following randomized mechanism achieves a revenue of 3.2, which is slightly more than that of the best deterministic mechanism:

$$a_A(v) = \begin{cases} 0 & v < 1.5 \\ \frac{4}{7} & v \in [1.5, 6) \\ \frac{6}{7} & v \in [6, 10) \\ 1 & v \geq 10 \end{cases} \qquad a_B(v) = \begin{cases} 0 & v < 1 \\ \frac{2}{7} & v \in [1, 3) \\ \frac{5}{7} & v \in [3, 8) \\ 1 & v \geq 8 \end{cases} \qquad a_C(v) = \begin{cases} 0 & v < 1 \\ \frac{2}{7} & v \in [1, 2) \\ \frac{4}{7} & v \in [2, 5) \\ \frac{5}{7} & v \in [5, 7) \\ \frac{6}{7} & v \in [7, 9) \\ 1 & v \geq 9 \end{cases}.$$

4.5 Key Concepts

In this section, we briefly overview of the key concepts of our work. Details are provided in Sections 4.7 and 4.8.

4.5.1 Bare Minimum Duality Preliminaries

4.5.1.1 Dual Terminology.

In this section, we introduce pictorial representations (Figures 4.7 and 4.8) of key aspects of a dual solution and define terminology relevant to the dual.

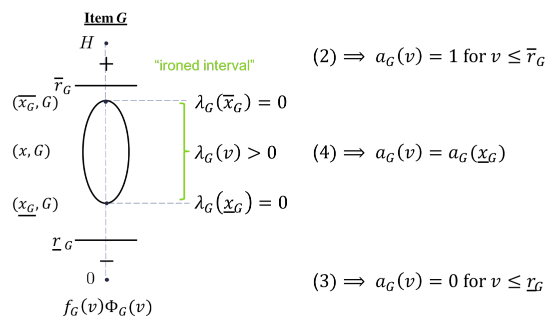


Figure 4.7: A pictorial interpretation of virtual values $f_G(v)\Phi_G^{\lambda;\alpha}(v)$ and the dual variable $\lambda_G(v)$, in addition to the concepts of endpoints of the zero region, ironing, an ironed interval, and the allocation in response.

The primal variables are $a_G(v)$ for all $G \in \mathcal{G}$, $v \in [0, H]$. Recall that we use $u_G(v) = \int_0^v a_G(w)dw$ to refer to the utility of (v, G). The dual variables are $\lambda_G(v)$, $\alpha_{G,G'}(v)$ for all $G, G' \in \mathcal{G}$, and $v \in [0, H]$. We first explain the role of these dual variables, and then describe the Lagrangian relaxation obtained using these dual variables.

Dual Variable λ. The λ dual variables correspond to incentive constraints between types of the same interest but different value. This dual controls *ironing*, as explained below. This really does correspond to ironing in the classical Myerson sense, only in value space.

An oval (as depicted in Figure 4.7) represents an *ironed interval*, a region where the dual variable $\lambda_G(\cdot)$ is non-zero.

- (Ironing) We say a type (v, G) is *ironed*, or that v is ironed in item G, if $\lambda_G(v) > 0$.
- (Ironed Intervals) For any type (x, G), the ironed interval containing x in G is defined by the bottom end point $\underline{x}_G = \sup\{v \leq x \mid \lambda_G(v) = 0\}$ and the top end point $\bar{x}_G = \inf\{v \geq x \mid \lambda_G(v) = 0\}$. Then for all $v \in (\underline{x}_G, \bar{x}_G)$, type (v, G) is ironed, $\bar{v}_G = \bar{x}_G$, and $\underline{v}_G = \underline{x}_G$.

As we will see later, dual best response (condition (4.4)) requires that if $\lambda_G(v) > 0$ then $a'_G(v) = 0$. In other words, the allocation rule a_G must be constant over ironed intervals. For any value x, an optimal allocation must satisfy that $a_G(x) = a_G(\underline{x}_G)$.

Dual Variable α. The α dual variables correspond to incentive constraints between types of the same value but different interest.

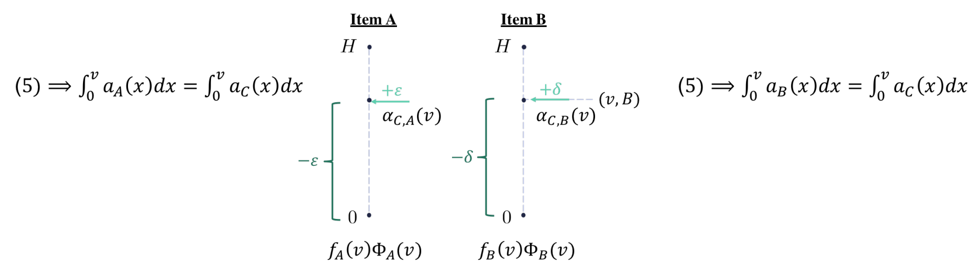


Figure 4.8: A pictorial representation of the dual variable α, in addition to the concepts of flow, preferable items, and equally preferable items. Flow is assumed to be coming from item C.

In Figure 4.8, a horizontal arrow into item A (or B) at v indicates that $\alpha_{C,A}(v)$ (or $\alpha_{C,B}(v)$) is non-zero. We write the following statements for $G \in \{A, B\}$.

- (Flow) We will call the value of $\alpha_{G',G}(v)$ the "flow into (v, G)" or the "flow into G at v." When we focus on the minimal partial-order example, we infer that flow into A or B comes from C in our figure.

Dual best response (condition (4.5)) requires that for $G \in \{A, B\}$, if $\alpha_{C,G}(v) > 0$ then $\int_0^v a_G(x)dx = \int_0^v a_C(x)dx$, or equivalently, $u_G(v) = u_C(v)$: a type with value v should have the same utility in C and G. Sending flow across interests forces the corresponding utilities to be the same.

Virtual Values. We will define a new variable, $\Phi^{\lambda,\alpha}(v)$ for all $v \in [0, H]$, and we will call the product $f(v)\Phi^{\lambda,\alpha}(v)$ the *virtual value*.[9] Once again, this is a generalization of Myerson's virtual value function to this more general setting.

[9]Whether we refer to Φ as the virtual value or Φf reflects whether we iron in the quantile space or the value space.

Figure 4.7 has a vertical axis ranging over values from 0 (at the bottom) to H (at the top), with a label of the item of focus G at the top. The point on the axis for any v represents the virtual value $f_G(v)\Phi_G^{\lambda,\alpha}(v)$.

Of particular interest to us is the region where the virtual value is 0 because this is the region (and the only region) for which a primal satisfying complementary slackness can have a randomized allocation. This is an interval if $(f_G\Phi_G^{\lambda,\alpha})(\cdot)$ is monotone in v (our solution ensures it is; details in Section 4.6.5).

- (Endpoints of Zero Region) We define the bottom end point of the zero virtual value region in G by $\underline{r}_G = \inf\{v \mid f_G(v)\Phi_G^{\lambda,\alpha}(v) \geq 0\}$ and the top end point $\bar{r}_G = \sup\{v \mid f_G(v)\Phi_G^{\lambda,\alpha}(v) \leq 0\}$.

In Figure 4.7 the horizontal black lines and signs indicate where the virtual values shift from positive sign to zero, $\overline{r}_G$, and from zero to negative sign, $\underline{r}_G$. Primal best response requires the allocation to satisfy $a_G(v) = 0$ for $v \leq \underline{r}_G$ (condition (4.2)) and $a_G(v) = 1$ for $v \geq \overline{r}_G$ (condition (4.3)).

4.5.1.2 The Lagrangian Dual.

The quality of a primal solution is measured by how well it solves the following Lagrangian relaxation induced by (λ, α). The quality of a dual solution is measured by the value of its induced Lagrangian relaxation. A dual is *better* if the value of its induced Lagrangian relaxation is *smaller*.

$$\begin{array}{ll} \text{Variables:} & a_G(v) \quad \forall G \in \mathcal{G},\ v \in [0, H] \\ \text{Maximize} & \displaystyle\sum_{G \in \mathcal{G}} \int_0^H f_G(v) \cdot a_G(v) \cdot \Phi_G^{\lambda,\alpha}(v) dv \\ \text{subject to} & a_G(v) \in [0, 1] \end{array}$$

$$\text{where} \qquad \varphi_G(v) = v - \frac{1 - F_G(v)}{f_G(v)} \qquad \text{and where} \qquad \Phi_G^{\lambda,\alpha}(v)$$

$$:= \varphi_G(v) + \frac{1}{f_G(v)} \left[-\lambda'_G(v) + \sum_{G' \in N^+(G)} \int_v^H \alpha_{G,G'}(w) dw - \sum_{G' : G \in N^+(G')} \int_v^H \alpha_{G',G}(w) dw \right]. \tag{4.1}$$

Before continuing, let's parse the Lagrangian relaxation. The only remaining constraints are that $a_G(v) \in [0, 1]$, and the objective is a linear function of these variables. This immediately implies that the solution to this LP relaxation will set $a_G(v) = 1$ whenever $\Phi_G^{\lambda;\alpha}(v) > 0$, and $a_G(v) = 0$ whenever $\Phi_G^{\lambda;\alpha}(v) < 0$. This implies that if there is any randomization, i.e., $a_G(v) \in (0, 1)$ then it must be that $\Phi_G^{\lambda;\alpha}(v) = 0$. The details of the definition of Φ are not so important here. (However, note that in the definition of Φ, the term λ' refers to the derivative of λ.)

4.5.1.3 Complementary Slackness.

Under strong duality, a (primal, dual) pair is optimal if and only if the primal and dual satisfy complementary slackness. In addition, if a dual (λ, α) is optimal, i.e. satisfies complementary slackness with some primal, then *any* primal is optimal if and only if it satisfies complementary slackness with (λ, α). Let's review complementary slackness in our setting. A primal a and dual (λ, α) satisfy complementary slackness if and only if:[10]

$$\text{(Primal best response)} \qquad \Phi_G^{\lambda;\alpha}(v) > 0 \quad \Rightarrow \quad a_G(v) = 1 \tag{4.2}$$

$$\Phi_G^{\lambda;\alpha}(v) < 0 \quad \Rightarrow \quad a_G(v) = 0 \tag{4.3}$$

$$\text{(Dual best response)} \qquad \lambda_G(v) > 0 \quad \Rightarrow \quad a'_G(v) = 0 \tag{4.4}$$

$$\alpha_{G,G'}(v) > 0 \quad \Rightarrow \quad \int_0^v a_G(x)dx - \int_0^v a_{G'}(x)dx = 0 \tag{4.5}$$

That is, a primal is a best response to a dual if all (v, G) with positive virtual value are awarded the item, and all (v, G) with negative virtual value are not. A dual is a best response to a primal if whenever a dual variable is non-zero, the corresponding local IC constraint is tight. The entire technical aspect of this chapter is using the constraints imposed by complementary slackness in (4.2-4.5) to reason about optimal mechanisms and their menu complexity.

[10] One can interpret these conditions as saying that the primal is an optimal solution to the Lagrangian relaxation, and the dual is the *worst* possible Lagrangian relaxation for the primal.

4.5.2 Menu Complexity is Unbounded: A Gadget and Candidate Instance

In this section, we provide a gadget that will be used in our menu complexity lower bound, and successively chain copies of it together to build our full construction. For one instance of our gadget, we provide a concrete potential dual, and prove that any allocation rule satisfying complementary slackness with it must have two distinct allocation probabilities. In order for this example to establish a lower bound of two, we must additionally:

- Establish that there exists a distribution D for which our dual is feasible. This is not covered in this section, and is deferred to our Master Theorem (Theorem 14).
- Establish that there exists an allocation rule which satisfies complementary slackness with this dual, thereby establishing that the dual is optimal (and any optimal allocation rule must satisfy complementary slackness with it). This is also not covered in this section, and is deferred to Section 4.8.

We begin below with our gadget, then successively chain copies together to establish a menu complexity lower bound of M for any $M > 0$. We recall the following facts established in the previous section:

1. A $+$ in any graphics at (x, G) represents a strictly positive Virtual Value, which implies that $a_A(x) = 1$ in any allocation rule satisfying CS. A $-$ in any graphics at (x, G) represents a strictly negative Virtual Value, which implies that $a_G(x) = 0$. (CS4.2-4.3)
2. A $\leftarrow$ in any graphics into A at x represents flow in. When there is flow into both A and B at the same point x, this implies that $u_A(x) = u_B(x)$. (CS4.5)
3. A point x in the middle of an oval in any graphics represents that x is contained in the interior of an ironed interval, and implies that $a(x) = a(y)$ where y is the bottom of the oval. (CS4.4)

4.5.2.1 Step One: the base gadget and a lower bound of $M = 2$.

Our base case example is depicted in Figure 4.9. We note each feature, and how it ties our hands with respect to the allocation rule via complementary slackness.

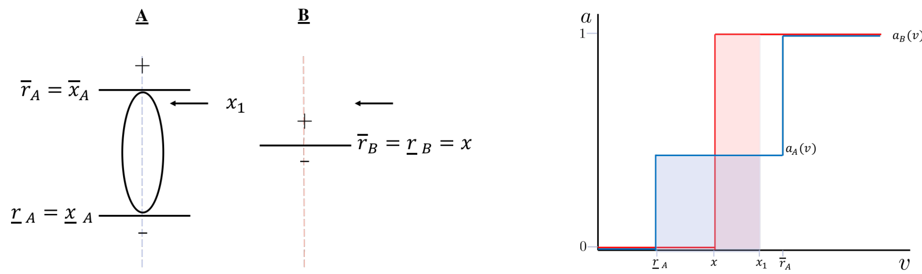


Figure 4.9: Left: Our first example that requires randomizing on A, containing an ironed interval $[\underline{r}_A, \overline{r}_A]$ (so $a_A(x_1) = a_A(\underline{r}_A)$) and flow into both A and B at x_1 (so $u_A(x_1) = u_B(x_1)$). Right: Primal best response dictates a price of x for item B, while A's allocation is 0 until $\underline{r}_A$ and 1 after $\overline{r}_A$. Equal preferability at x_1 forces $u_A(x_1)$ (the red area) equals $u_B(x_1)$ (the blue area); the ironed interval $[\underline{r}_A, \overline{r}_A]$ requires $a_A(\cdot)$ to be constant in this region, hence we must have $a_A(\underline{r}_A) \in (0, 1)$.

- In item B, there is a single point $x < \overline{r}_A$ for which $f_B(x)\Phi_B^{\lambda,\alpha}(x) = 0$. That is, $\overline{r}_B = \underline{r}_B = x$. Then condition (4.2) implies that $a_B(v) = 1$ for $v > x$.

- There is flow into both items A and B at $x_1 > x$. That is, $\alpha_{C,A}(x_1), \alpha_{C,B}(x_1) > 0$. Condition (4.5) implies that A and B must be equally preferable at x_1, that is, $\int_0^{x_1} a_A(w)dw = \int_0^{x_1} a_B(w)dw$. Note that $a_B(w) > 0$ for $w \in (x, x_1]$, hence $\int_0^{x_1} a_B(w)dw > 0$. Then to have $\int_0^{x_1} a_A(w)dw > 0$, because $a_A(\cdot)$ is monotone, it must be the case that $a_A(x_1) > 0$.

- The point x_1 has $f_A(x_1)\Phi_A^{\lambda,\alpha}(x_1) = 0$ and is in an ironed interval $[\underline{r}_A, \overline{r}_A]$ where $\underline{r}_A < x$, that is, this ironed interval is the entire region of values that have virtual value zero in item A and it contains both x_1 and x. Because x_1 is in an ironed interval in A, then the allocation is constant, so $a_A(\underline{r}_A) = a_A(x_1)$, which we have already established must be positive.

- For whatever value that $a_A(\underline{r}_A)$ takes on, because $\underline{r}_A < x$, to satisfy equal preferability at x_1 (again, that $\int_0^{x_1} a_A(w)dw = \int_0^{x_1} a_B(w)dw$), we must have $a_B(x) > a_A(\underline{r}_A)(> 0)$, resulting in at least two distinct non-zero probabilities of allocation.

To complete the example, (1) there is no other flow: for all $v \neq x_1$, $\alpha_{C,A}(v) = \alpha_{C,B}(v) = 0$, and (2) item C is unironed everywhere: $\lambda_C(v) = 0$ for all v. This base gadget forces randomization for the allocation of item A because the utility of x_1

must be equal at A and B, but the allocation of item B must be zero below x, while the allocation of item A must be non-zero.

4.5.2.2 Step Two: two chains and a lower bound of $M = 3$.

Our second example (see Figure 4.10) contains the relevant features from the first example, but extends it to add an additional constraint: we replace the condition $\overline{r}_B = \underline{r}_B = x$ with an ironed interval $[\underline{r}_B, \overline{r}_B]$ where $\underline{r}_B < \underline{r}_A < \overline{r}_B < \overline{r}_A$. We claim that this example requires us to randomize on both items. Intuitively, this is because we now have two constraints on utilities that must be satisfied, so two degrees of freedom seems necessary.

- There is flow into both items A and B at $x_1 \in (\overline{r}_B, \overline{r}_A)$: $\alpha_{C,A}(x_1), \alpha_{C,B}(x_1) > 0$. Condition (4.2) implies that $a_B(v) = 1$ for $v > \overline{r}_B$, so to satisfy equal preferability, we must have $a_A(x_1) > 0$.

- The point x_1 has $f_A(x_1)\Phi_A^{\lambda,\alpha}(x_1) = 0$ and is in an ironed interval $[\underline{r}_A, \overline{r}_A]$ where $\underline{r}_A < \overline{r}_B$. As x_1 is in an ironed interval in A, then the allocation is constant, so $a_A(\underline{r}_A) = a_A(x_1) > 0$.

- There is flow into both items A and B at $x_2 \in (\underline{r}_A, \overline{r}_B)$: $\alpha_{C,A}(x_2), \alpha_{C,B}(x_2) > 0$. Since $a_A(x_2) > 0$—it lies in the ironed interval in A, so $a_A(x_2) = a_A(\underline{r}_A)$—then to satisfy equal preferability at x_2, we must have $a_B(x_2) > 0$.

- The point x_2 has $f_B(x_2)\Phi_B^{\lambda,\alpha}(x_2) = 0$ and is in an ironed interval $[\underline{r}_B, \overline{r}_B]$ where $\underline{r}_B < \underline{r}_A$. As x_2 is in an ironed interval in B, then the allocation is constant, so $a_B(\underline{r}_B) = a_B(x_2) > 0$.

- For whatever value that $a_B(\underline{r}_B)$ takes on, because $\underline{r}_B < \underline{r}_A$, then to satisfy equal preferability at x_2 (that $\int_0^{x_2} a_A(w)dw = \int_0^{x_2} a_B(w)dw$), we must have $a_A(\underline{r}_A) > a_A(\underline{r}_B)(> 0)$.

- For whatever value that $a_A(\underline{r}_A)$ takes on, because $\underline{r}_A < \overline{r}_B$, then to satisfy equal preferability at x_1 ($\int_0^{x_1} a_A(w)dw = \int_0^{x_1} a_B(w)dw$), we must have $a_B(\overline{r}_B) > a_A(\underline{r}_A)(> a_A(\underline{r}_B) > 0)$, resulting in at least *three* distinct non-zero probabilities of allocation.

Again, (1) there is no other flow: for all $v \neq x_1, x_2$, $\alpha_{C,A}(v) = \alpha_{C,B}(v) = 0$, and (2) item C is unironed everywhere: $\lambda_C(v) = 0$ for all v.

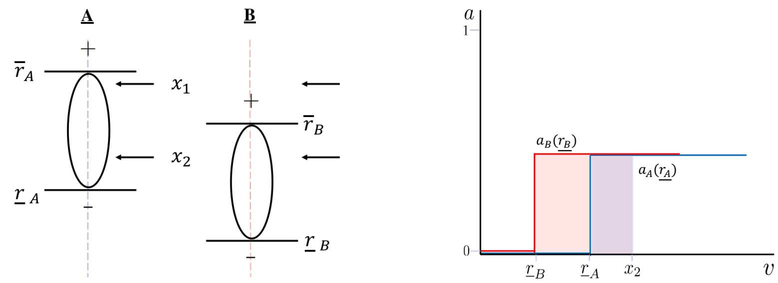


Figure 4.10: Left: Our second example, which requires randomization on both A and B.
Right: If $a_A(\underline{r}_A) \leq a_B(\underline{r}_B)$, then $u_A(x_2) < u_B(x_2)$ (the blue region is smaller than the red), which violates complementary slackness.

Observe that in both examples, we reason from where we have one item with positive virtual value and the other with virtual value zero downward that, in order to satisfy a number of equal preferability constraints, because ironed intervals force the allocation to be constant, then at every point, the allocation must be non-zero. Then, we reason upward that, because the ironed intervals are interleaving between the items and never aligned, the allocation must strictly increase at each point of interest in order to satisfy equal preferability. This is precisely the reasoning we will use to construct and prove an arbitrarily large instance and menu.

4.5.2.3 Step Three: four chains and a lower bound of $M = 5$.

In this section, we take one more step towards our general construction. The first example presents our base gadget, and the second example chains two copies together. In this section, we simply confirm how the gadgets interact as we chain more and more together, bouncing back and forth from A to B.

Nonzero allocation probabilities. First, we see that the allocation at every ironed value v such that $\Phi_G^{\lambda,\alpha}(v) = 0$ must be nonzero: $a_G(v) > 0$. The argument holds for each of $(x_1, A), (x_2, B)$, (x_3, A), and (x_4, B). Below we iterate the same argument made in the two previous sections, skipping some details.

- Suppose $a_B(x_1) > 0$ (this is necessarily true by Fact 1), and thus $u_B(x_1) > 0$.
- By Fact 2, $u_A(x_1) = u_B(x_1) > 0$. Then $a_A(x_1) > 0$.

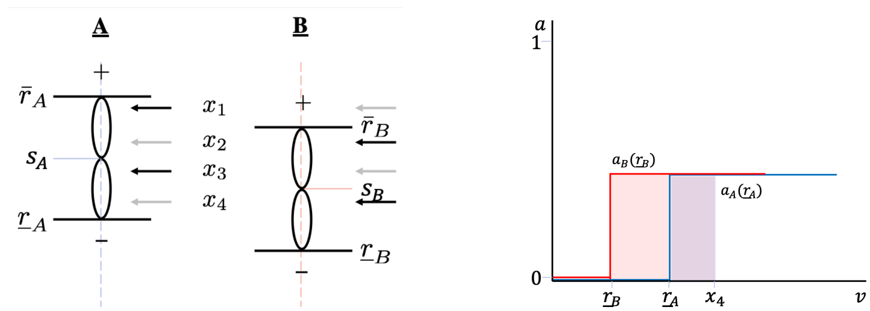


Figure 4.11: Left: An optimal dual for our example distributions, which will require at least 4 distinct allocation probabilities.
Right: If $a_A(\underline{r}_A) \leq a_B(\underline{r}_B)$, then $u_A(x_2) < u_B(x_2)$ (the blue region is smaller than the red), which violates complementary slackness via Fact 2 at x_4.

- By Fact 3, $a_A(s_A) = a_A(x_1) > 0$. This also implies that $u_A(x_2) > 0$.
- Now, again by Fact 2, $u_B(x_2) = u_A(x_2) > 0$, so $a_B(x_2) > 0$.
- Now, again by Fact 3, $a_B(x_3) = a_B(x_2) > 0$, so $u_B(x_3) > 0$.
- Again by Fact 2, $u_A(x_3) = u_B(x_3) > 0$, so $a_A(x_3) > 0$.
- By Fact 3, $a_A(x_4) = a_A(x_3) > 0$, so $u_A(x_4) > 0$.
- Finally by Fact 2, $u_B(x_4) = u_A(x_4) > 0$.

Essentially, if any of these allocations must be positive, it forces the rest of them, working downwards, to be positive. And, by Fact 1, $a_B(x_1) = 1$, so $u_B(x_1) > 0$. Hence the rest of the implications follow, so the allocation must be nonzero throughout this region.

Distinct allocation probabilities. Now, given that the allocation must be nonzero at every point in this range, we argue that it must be distinct at all of the points of interest. Fix some nonzero $a_B(\underline{r}_B)$, and note by Fact 3 that $a_B(v) = a_B(\underline{r}_B)$ for all $v \in [\underline{r}_B, s_B]$. By Fact 1, $a_G(v) = 0$ for $v < \underline{r}_G$. Because $\underline{r}_B < \underline{r}_A$, then to have $u_A(x_4) = u_B(x_4)$, since $u_A(x_4) = \int_{\underline{r}_A}^{x_4} a_A(w)dw = (x_4 - \underline{r}_A)a_A(\underline{r}_A)$ and $u_B(x_4) = \int_{\underline{r}_B}^{x_4} a_B(w)dw = (x_4 - \underline{r}_B)a_B(\underline{r}_B)$, then we must have a distinct $a_A(\underline{r}_A) > a_B(\underline{r}_B)$. This is depicted on the right side in Figure 4.11. Then, by Fact 3, $a_A(x_3) = a_A(\underline{r}_A) > a_B(\underline{r}_B)$.

The argument extends inductively for $(x_3, B), (x_2, A)$, and (x_1, B): we show it with (x_3, B). Note that $u_A(x_4) = u_B(x_4)$ and suppose the inductive hypothesis of

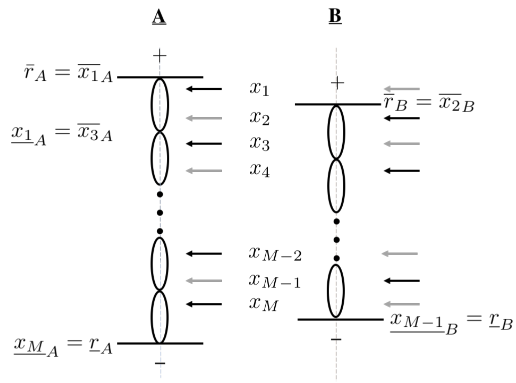


Figure 4.12: Our candidate dual instance: a top chain that spans the entire region of zero virtual values for both A and B with no gaps between the ironed intervals that comprise the chain. There is flow into A and B at every point x_i in the chain.

$a_A(x_4) > a_B(x_4)$, where $a_A(x_3) = a_A(x_4)$ and $a_B(x_4) = a_B(\underline{r}_B)$ by Fact 3. Hence $u_A(s_B) > u_B(s_B)$. Then in order to have $u_A(x_3) = u_B(x_3)$, we must have $a_B(x_3) > a_A(x_3)$.

The result is four distinct allocation probabilities in these four regions, and five in total (including the deterministic option to get the item w.p. one). Essentially, this example only has two ironed intervals in A and B each with four points of interest. Our full construction below lets the number of ironed intervals grow with M.

4.5.2.4 Final Step: M chains and a lower bound of M.

It is possible to extend the examples above by continuing to interleave ironed intervals with flow coming in. The combination of the equal preferability constraints and the inability to increase the allocation in the middle of an ironed interval is what requires us to randomize differently within each interval, forcing any number of menu options. Details are given in Section 4.7, where we formally define this "top chain" structure (Definition 9) and construct the candidate dual instance, which is depicted in Figure 4.12. For example, our first example has a top chain of length one, the second of length two, and the third of length four. Theorem 4 proves that there exists a primal instance that satisfies complementary slackness with the defined dual. This proves both that our dual is optimal, and thus *any* optimal primal must satisfy complementary slackness with it, giving us Theorem 3.

Theorem 3. *Mechanisms that satisfy complementary slackness with a dual containing a top chain of length M have menu complexity at least M. Moreover, for all M, there exists a distribution D over three partially-ordered items for which a dual with top chain of length M is feasible.*

The "Moreover, . . . " part of the theorem is due to our Master Theorem (Theorem 14). The formal statement is a bit technical, and can be found in Section 4.11.

4.5.3 Menu Complexity is Finite: Brief Highlight

In Section 4.8, we discuss our approach for characterizing the optimal mechanism for our 3-item minimal instance. We prove essentially that the interleaving of ironed intervals used in the construction of the previous section is the worst case (in terms of menu complexity). We do this by specifying a subclass of optimal duals (that we call *best duals*) using two new dual operations, *double swaps* and *upper swaps*. We then leverage the structure of the best duals to give an algorithm that recovers the optimal primal from any best dual, and prove that the resulting mechanism has finite menu complexity.

Theorem 4. *For any best dual solution, the primal recovery algorithm returns a primal with finite menu complexity that satisfies complementary slackness (and is therefore optimal).*

We conclude with one vignette regarding how the menu complexity can be unbounded but not infinite. Two crucial aspects of the "top chain" structure from our examples (generalized in Figure 4.12) are that: (1) the ironed intervals for A and B are interleaving—this is what "keeps the chain going" and (2) the sequences for A and B terminate at *different* bottom endpoints. The latter is a bit subtle, but the idea is that if the two chains terminate at the same bottom endpoint v, then this entire process can be aborted and simply setting v as the reserve for all items satisfies complementary slackness. So while in principle, this top chain structure could indeed be countably infinite, it cannot also satisfy (1) and (2). This is because the monotone convergence theorem states that both chains do indeed converge to some bottom endpoint, and interleaving then guarantees that this bottom endpoint must be the same.

4.5.4 One Last Example

In this section, we construct an example by applying the Master Theorem (Theorem 14) to the dual in Figure 4.11. The customer prior distribution in the example consists of the marginal distributions depicted in Figures 4.5 and 4.6. The distributions for A and B do not satisfy DMR, and, using the ideas from the previous subsections, we will see that the optimal mechanism is randomized.

We can use the revenue curve procedure from Section 4.4 to determine the optimal pricing for this example. It produces the curves in Figure 4.6, telling us that the optimal price to set on item C is 8, which will result in prices of 9 on item A and 8 on item B. This gives $R_{ABC}(8) = 3.155$. However, as we have seen in Subsection 4.5.2, for the dual in Figure 4.11 (which corresponds to this distribution) to satisfy complementary slackness with a mechanism, the mechanism must have a good deal of randomization.

In Subsection 4.5.2, we reasoned that the allocation probability must be distinct at each of the points (x_1, A), (x_2, B), (x_3, A), and (x_4, B). We also saw that if we fixed the allocation at (x_4, B), there was only one way to satisfy the rest of the complementary slackness constraints, forming a system of equations. The primal recovery algorithm described in the proof of Theorem 4 goes through solving this system of equations, ensuring that any other additional complementary slackness constraints are met, and that no pathological structures that might prevent a solution from existing can arise. Applying this algorithm to our example results in the following optimal randomized mechanism:

$$a_A(v) = \begin{cases} 0 & v < 1.5 \\ \frac{4}{7} & v \in [1.5, 6) \\ \frac{6}{7} & v \in [6, 10) \\ 1 & v \geq 10 \end{cases} \qquad a_B(v) = \begin{cases} 0 & v < 1 \\ \frac{2}{7} & v \in [1, 3) \\ \frac{5}{7} & v \in [3, 8) \\ 1 & v \geq 8 \end{cases} \qquad a_C(v) = \begin{cases} 0 & v < 1 \\ \frac{2}{7} & v \in [1, 2) \\ \frac{4}{7} & v \in [2, 5) \\ \frac{5}{7} & v \in [5, 7) \\ \frac{6}{7} & v \in [7, 9) \\ 1 & v \geq 9 \end{cases}.$$

The mechanism achieves a revenue of 3.2, which is slightly more than that of the best deterministic mechanism.

4.6 Full Preliminaries

While this chapter focuses on the three-item case, it's illustrative (and perhaps cleaner) to provide notation for general partially-ordered items. In general, there are m partially-ordered items. Item G can be better than, worse than, or incomparable to item G', and we'll use the relation $G \succ G'$ to denote that G is better than G'. We refer to the set of items as $\mathcal{G}$, and use $N^+(G)$ to denote the set of items $G' \in \mathcal{G}$ for which $G' \succ G$, but there is no G'' with $G' \succ G'' \succ G$ (i.e. the items "immediately better" than G, or the 1-out-neighborhood of G in a graphic representation). There is a single buyer with a (value, interest) pair (v, G), who receives value v if they are awarded an item $\succeq G$. An instance of the problem consists of a joint probability distribution over $[0, H] \times \mathcal{G}$, where H is the maximum possible value of any bidder for any item. We will use f to denote the density of this joint distribution, with $f_G(v)$ denoting the density at (v, G). We will also use $F_G(v)$ to denote $\int_0^v f_G(w)dw$, and q_G to denote the probability that the bidder's interest is G. Note that $F_G(H) = q_G < 1$, so $F_G(\cdot)$ is not the CDF of a distribution (although $F_G(\cdot)/q_G$ is the CDF of the marginal distribution of v conditioned on interest G).

We'll consider (w.l.o.g.) direct truthful mechanisms, where the bidder reports a (value, interest) pair and is awarded a (possibly randomized) item. For a direct mechanism, we'll define $a_G(v)$ to be the probability that item G is awarded to a bidder who reports (v, G), and $p_G(v)$ to be the expected payment charged. Then a buyer's utility for reporting any (v', G') where G' doesn't dominate G is $-p_{G'}(v')$, and the utility for reporting any (v', G') where G' dominates G is $v \cdot a_{G'}(v') - p_{G'}(v')$.

At this point, one can write a primal LP that maximizes expected revenue subject to incentive constraints, manipulate the LP, and consider a Lagrangian relaxation (and all of this is done in Fiat et al. [2016]; Devanur and Weinberg [2017]).

4.6.1 Formulating the Optimization Problem

The "default" way to write the continuous LP characterizing the optimal mechanism would be to maximize $\sum_{G \in \mathcal{G}} \int_0^H f_G(v)p_G(v)dv$ (the expected revenue) such that everyone prefers to tell the truth than to report any other type. As observed in Fiat et al. [2016], it is without loss of generality to only consider mechanisms that award bidders their declared item of interest with probability in $[0, 1]$, and all other items with probability 0.[11] Also observed in Fiat et al. [2016] is that Myerson's payment identity holds in this setting as well, and any truthful mechanism must satisfy $p_G(v) = va_G(v) - \int_0^v a_G(w)dw$ (this also implies that the bidder's utility when truthfully reporting (v, G) is $u_G(v) = \int_0^v a_G(w)dw$). This allows us to drop the payment variables, and follow Myerson's analysis to recover:[12]

$$\mathbb{E}[\text{revenue}] = \sum_{G \in \mathcal{G}} \int_0^H f_G(v) \cdot p_G(v)dv = \sum_{G \in \mathcal{G}} \int_0^H f_G(v)a_G(v) \left(v - \frac{1 - F_G(v)}{f_G(v)} \right) dv$$

The experienced reader will notice that $v - \frac{1-F_G(v)}{f_G(v)}$ is exactly Myerson's virtual value for the conditional distribution $F_G(\cdot)/q_G$, which we'll denote by $\varphi_G(v)$. At this point, we still have a continuous LP with only allocation variables, but still lots of truthfulness constraints. Fiat et al. [2016] observe that many of these constraints are redundant, and in fact it suffices to only make sure that when the bidder has (value, interest) pair (v, G) they:

- Prefer to tell the truth rather than report any other (v', G). This is accomplished by constraining $a_G(\cdot)$ to be monotone non-decreasing (exactly as in the single-item setting).
- Prefer to tell the truth rather than report any other $(v, G' \in N^+(G))$. This is accomplished by constraining $\int_0^v a_G(w)dw \geq \int_0^v a_{G'}(w)dw$ (as the LHS denotes the utility of the buyer for reporting (v, G) and the RHS denote the utility of the buyer for reporting (v, G')).

All of these constraints together imply that (v, G) also does not prefer to report any other (v', G').[13] Below, we will now formulate the Primal LP and its Lagrangian

[11] To see this, observe that the bidder is just as happy to get nothing instead of an item that doesn't dominate their interest. See also that they are just as happy to get their interest item instead of any item that dominates it. It will also make this option no more attractive to any bidder considering misreporting. So starting from a truthful mechanism, modifying it to only award the item of declared interest or nothing cannot possibly violate truthfulness.

[12] For the familiar reader, this derivation is routine, so we omit it. The unfamiliar reader can refer to [Myerson, 1981; Hartline, 2013] for this derivation.

[13] For example, if (v, G) prefers truthful reporting to reporting (v, G') where $G' \succ G$, and (v, G') prefers truthful reporting to reporting (v', G'), then since (v, G) gets the same utility for reporting (v, G') as type (v, G') does for truthfully reporting, (v, G) prefers truthful reporting to reporting (v', G').

relaxation. This derivation is not a new result, but important to understanding our approach. So we'll go through some of the steps to help provide some intuition for the reader, but omit any calculations and proofs.

4.6.2 The Primal

With the above discussion in mind, we can now formulate our primal continuous LP.

$$\begin{array}{ll} \text{Variables:} & a_G(v),\ \forall G \in \mathcal{G},\ v \in [0, H] \\ \text{Maximize} & \displaystyle\sum_{G \in \mathcal{G}} \int_0^H f_G(v) a_G(v) \varphi_G(v) dv \end{array}$$

$$\begin{array}{lll} \text{subject to} & a'_G(v) \geq 0 & \forall G \in \mathcal{G}\ \forall v \in [0, H] \text{ (dual variables } \lambda_G(v) \geq 0) \\ & \displaystyle\int_0^v a_G(x)dx - \int_0^v a_{G'}(x)dx \geq 0 & \forall G \in \mathcal{G},\ G' \in N^+(G)\ \forall v \in [0, H] \text{ (dual vars } \alpha_{G,G'}(v) \geq \\ & a_G(v) \in [0, 1] & \forall G \in \mathcal{G},\ \forall v \in [0, H] \text{ (no dual variables)} \end{array}$$

The first constraint requires that $a_G(\cdot)$ is monotone non-decreasing for all G. If an allocation rule is not monotone, it cannot possibly be part of a truthful mechanism. As discussed above, Myerson's payment identity combined with monotonicity guarantees that (v, G) will always prefer to report (v, G) instead of (v', G). The second constraint directly requires that the utility of (v, G) for reporting (v, G) is at least as high as for reporting (v, G') (also discussed above). The final constraint simply ensures that the allocation probabilities lie in $[0, 1]$.

4.6.3 Derivation of the Partial Lagrangian Dual

Moving the first two types of constraints from the primal to the objective function with multipliers $\lambda_G(v)$ and $\alpha_{G,G'}(v)$ respectively gives the partial Lagrangian primal:

$$\max_{a: a_G(v) \in [0,1]\ \forall G \in \mathcal{G}, \forall v \in [0,H]}\ \min_{\lambda, \alpha \geq 0} \mathcal{L}(a; \lambda, \alpha)$$

where

$$\mathcal{L}(a;\lambda,\alpha) :=$$
$$\sum_{G\in\mathcal{G}} \int_0^H \left[f_G(v)a_G(v)\varphi_G(v) + \sum_{G'\in N^+(G)} \alpha_{G,G'}(v)\cdot\left[\int_0^v a_G(x)dx - \int_0^v a_{G'}(x)dx\right] + \lambda_G(v)a'_G(v)\right] dv.$$

This gives the corresponding partial Lagrangian dual of

$$\min_{\lambda,\alpha\geq 0} \max_{a:a_G(v)\in[0,1]\,\forall G\in\mathcal{G},\forall v\in[0,H]} \mathcal{L}(a;\lambda,\alpha).$$

Note however that we can rewrite $\mathcal{L}(a;\lambda,\alpha)$ by using integration by parts on the $a'_G(v)$ term to get $a_G(v)$ terms, using that $a_G(0)=0$ and $\lambda_G(H)=0$ without loss:

$$\int_0^H \lambda_G(v)a'_G(v)dv = \lambda_G(v)a_G(v)\mid_0^H - \int_0^H \lambda'_G(v)a_G(v)dv = -\int_0^H \lambda'_G(v)a_G(v)dv$$

As in [FGKK '16], this uses the facts that $\lambda_G(\cdot)$ is continuous and equal to 0 at any point that $a'_G(v)=\infty$, which occurs at only countably many points. Then, collecting the $a_G(v)$ terms gives:

$$\begin{aligned}\mathcal{L}(a;\lambda,\alpha) &= \sum_{G\in\mathcal{G}} \int_0^H \Bigg[f_G(v)a_G(v)\varphi_G(v) \\ &\quad + \sum_{G'\in N^+(G)} \alpha_{G,G'}(v)\cdot\left[\int_0^v a_G(x)dx - \int_0^v a_{G'}(x)dx\right] - \lambda'_G(v)a_G(v)\Bigg] dv \\ &= \sum_{G\in\mathcal{G}} \int_0^H f_G(v)a_G(v)\Phi_G^{\lambda,\alpha}(v)dv\end{aligned}$$

where we define

$$\Phi_G^{\lambda,\alpha}(v) := \varphi_G(v) + \frac{1}{f_G(v)}\cdot\left[\sum_{G'\in N^+(G)} \int_v^H \alpha_{G,G'}(x)dx - \sum_{G':G\in N^+(G')} \int_v^H \alpha_{G',G}(v)dx\right] - \frac{1}{f_G(v)}\lambda'_G(v).$$

Then we can write that the Lagrangian dual problem is

$$\min_{\lambda,\alpha\geq 0} \max_{a:a_G(v)\in[0,1]\,\forall G\in\mathcal{G},\forall v\in[0,H]} \sum_{G\in\mathcal{G}} \int_0^H f_G(v)a_G(v)\Phi_G^{\lambda,\alpha}(v)dv.$$

4.6.4 More Dual Terminology

Minimal dual terminology is first introduced in Section 4.5.1. Here, we add a few additional terms.

Dual best response (condition (4.5)) implies the following.

- (Preferable Items) To satisfy complementary slackness, for any x such that $\alpha_{G,G'}(x) > 0$, we must have $u_{G'}(x) \geq u_{G''}(x) \quad \forall G'' \in N^+(G)$. This is because (a) $u_G(x) = u_{G'}(x)$ by complementary slackness and (b) $u_G(x) \geq u_{G''}(x) \quad \forall G'' \in N^+(G)$ by incentive compatibility.

- (Equally Preferable Items) By the above, to satisfy complementary slackness with any dual with $\alpha_{G,G'}(x) > 0$ and $\alpha_{G,G''}(x) > 0$, we must have $u_{G'}(x) = u_{G''}(x)$.

4.6.5 Review of Dual Properties

- (Rerouting Flow Among $N^+(G)$) If $G', G'' \in N^+(G)$ and we decrease $\alpha_{G,G'}(v)$ by ε and increase $\alpha_{G,G''}(v)$ by ε, then $v' \leq v$, $f_{G'}(v')\Phi_{G'}^{\lambda,\alpha}(v')$ decreases by ε and $f_{G''}(v')\Phi_{G''}^{\lambda,\alpha}(v')$ increases by ε. All other virtual values, including all of those within G, remain the same.

- (Utility based on the dual) We can often simplify how utility is written in terms of the dual and complementary slackness constraints. If $\underline{x}_G < x < y < \overline{x}_G$, then allocation in ironed intervals implies $u_G(y) = u_G(x) + a_G(y)(y - x)$.

- (Allocation to Nonzero Virtual Values) As shown above in Subection 4.5.1.2, the dual variables (1) determine the virtual welfare functions $\Phi^{\lambda,\alpha}(\cdot)$ and (2) are chosen to minimize the maximum virtual welfare under $\Phi^{\lambda,\alpha}(\cdot)$. For an optimal dual solution, the optimal mechanism will simply be the corresponding virtual welfare maximizer that satisfies complementary slackness. Parts of this mechanism are easy to predict if the virtual value functions are sign-monotone, which we will later ensure that they are. Assuming this, we can talk about the virtual values in terms of three regions: positives, negatives, and zeroes.

- (Ironing and Proper Monotonicity.) We say that a dual satisfies *proper monotonicity* if $f_G \cdot \Phi_G^{\lambda,\alpha}(\cdot)$ is monotone non-decreasing (note the multiplier of f_G).

As shown in [FGKK '16; DW '17], for all α, there exists a λ such that (λ, α) is properly monotone.

- (Boosting can only improve the dual.) Given any dual with properly monotone virtual values, if there exists v such that $f_G(v)\Phi_G^{\lambda;\alpha}(v) < 0$, then for any $G' \in N^+(G)$, incrementing $\alpha_{G,G'}(v)$ by $f_G(v)\Phi_G^{\lambda;\alpha}(v)$ only improves the dual. By proper monotonicity, for all $v' \leq v$, $f_G(v')\Phi_G^{\lambda;\alpha}(v') < f_G(v)\Phi_G^{\lambda;\alpha}(v) < 0$, hence increasing $\alpha_{G,G'}(v)$ will not create any positives within G, not hurting the dual objective. Sending flow into an item G' can only help by making positives less so, and does not increase any virtual values (but it's possible that it doesn't strictly help). This operation is coined *boosting* in [DW '17]. While it is clear that G should send the flow, the remaining question is *which* $G' \in N^+(G)$ should the flow be sent to. This is the bulk of our analysis.

- By sign monotonicity, $v > \bar{r}_G$ has a positive virtual value, and thus the allocation rule must set $a_G(v) = 1$, otherwise it is not maximizing virtual welfare.

- Similarly, for values with negative virtual values, that is, $v < \underline{r}_G$, it must be that $a_G(v) = 0$.

From these observations, we can conclude that the flow out of C is identical to the flow out of the root node (day n) in the FedEx solution. That is,

$$\alpha_{C,A}(v) + \alpha_{C,B}(v) = \begin{cases} 0 & v > \bar{r}_C \\ -\hat{R}_C''(v)/f_C(v) & v \leq \bar{r}_C. \end{cases}$$

where $R_C(\cdot)$ is defined as in Definition 5, $\hat{R}_C(\cdot)$ is the least concave upper bound on $R_C(\cdot)$, and $\hat{R}_C''(\cdot)$ is the second derivative of this function with respect to v.

We conclude with a fundamental result from [FGKK '16].

Theorem 5 (Proper Ironing [FGKK '16]). *Given all dual variables α, suppose $\lambda_G(v) = 0$ for all (v, G). Then $f_G(v)\Phi_G^{\lambda;\alpha}(v)$ is defined for all (v, G). We define $\Gamma_G(v) = -\int_0^v f_G(x)\Phi_G^{\lambda,\alpha}(x)dx$, and $\hat{\Gamma}_G(\cdot)$ is the least concave upper bound on this function. Then setting $\lambda_G(v) = \hat{\Gamma}_G(v) - \Gamma_G(v)$ defines a continuous and differentiable $\lambda_G(\cdot)$ that, with the update of $\Phi_G^{\lambda;\alpha}(\cdot)$ based on $\lambda_G(\cdot)$, results in the proper monotonicity of $f_G(\cdot)\Phi_G^{\lambda;\alpha}(\cdot)$.*

4.7 Formal Construction of a Candidate Dual Instance

We extend the above examples from SubSubsection 4.5.2 to require any number of menu options. As in the two examples, we can reason from the top downward that the allocation at the bottom of every ironed interval must be positive, and reason from the bottom upward that the allocation must strictly increase for each new overlapping ironed interval we encounter, yielding all different menu options. We formally define this interleaving structure and call it a "chain," depicted in Figure 4.13. As another sanity check: each new point in the chain induces a new equality that has to be satisfied. So if the chain is of length M, intuition suggests that we should need M degrees of freedom to possibly satisfy complementary slackness (but this is just intuition).

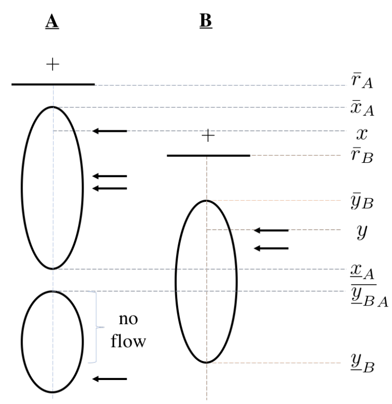


Figure 4.13: This is an example of a chain that consists of the points $\{(x, A), (y, B)\}$. It is a top chain as $x > \bar{r}_B$. Note that (y, B) is preceded by (x, A) as there is flow into B at y and $y > \underline{x}_A$, and there is no flow into B for any $v \in (y, \bar{y}_B]$. The chain terminates at (y, B) since there is no flow into A for any $v \in [\underline{y}_B, \overline{\underline{y}_B}_A]$.

Definition 9 (Top chain)**.** A sequence $(x_1, A), (x_2, B), (x_3, A), \cdots$ of points that switch between items A and B is called a *chain* if the following hold:

- $\Phi_A^{\lambda,\alpha}(x) = 0$ for all (x, A) in the chain and $\Phi_B^{\lambda,\alpha}(y) = 0$ for all (y, B) in the chain.
- $\alpha_{C,A}(x) > 0$ for all (x, A) in the chain and $\alpha_{C,B}(y) > 0$ for all (y, B) in the chain.
- $\lambda_A(x) > 0$ for all (x, A) in the chain and $\lambda_B(y) > 0$ for all (y, B) in the chain.

- $\underline{x_i}_A < x_{i+1} < x_i$ if (x_i, A) is in the chain and $\underline{x_i}_B < x_{i+1} < x_i$ if (x_i, B) is in the chain.

We call a chain the *top chain* if $x_1 > \overline{r}_B$.

Note that if any of these conditions do not hold, the mechanism has an easier solution. If any point v in the zero regions of both A and B were unironed, we could just set a price of v for both. If the chains did not interleave with flow alternating in, our series of constraints would end. The top chain structure (and it is key that it is a top chain) prevents this.

We now provide a complete proof of Theorem 3. First, we provide a construction of our candidate dual, which is depicted in Figure 4.12. The instance uses definition 9 of a top chain.

Construction of candidate dual instance:

- Let there exist no point at which A and B both have virtual value zero and both are unironed, that is, there is no v such that $\Phi_A^{\lambda,\alpha}(v) = \Phi_B^{\lambda,\alpha}(v) = 0$ and $\lambda_A(v) = \lambda_B(v) = 0$.
- Let $\overline{r}_A > x_1 > \overline{r}_B > x_2 > x_3 > \cdots > x_M > \underline{r}_B > \underline{r}_A$. The dual has a top chain of length M defined by $(x_1, A), (x_2, B), \ldots, (x_M, A)$.
- In addition, we have flow into the other item at each point in the chain: let $\alpha_{C,B}(x_i) > 0$ for all (x_i, A) in the chain as well as $\alpha_{C,A}(x_i) > 0$ for all (x_i, B) in the chain.
- Let $\lambda_C(v) = 0$ for all v, *i.e.*, item C is unironed everywhere.
- For all v where α has not already been defined, let $\alpha_{C,A}(v) = \alpha_{C,B}(v) = 0$.

We first make some remarks that follow directly from our construction. All the remarks below (only) talk about our dual and any feasible primal that satisfies complementary slackness with our dual.

Remark 6. *For all* $i \in \{1, 3, \cdots, M-2\}$*, we have* $x_i, x_{i+1} \in [\underline{x_i}_A, \overline{x_i}_A] = [\underline{x_{i+1}}_A, \overline{x_{i+1}}_A]$*. Since this interval is ironed, we have* $\lambda_A(v) > 0 \implies a'_A(v) = 0$ *for* v *in this interval. Thus,* $a_A(x_i) = a_A(x_{i+1})$.

Remark 7. *For all* $i \in \{2, 4, \cdots, M-1\}$*, we have* $x_i, x_{i+1} \in [\underline{x_i}_B, \overline{x_i}_B] = [\underline{x_{i+1}}_B, \overline{x_{i+1}}_B]$*. Since this interval is ironed, we have* $\lambda_B(v) > 0 \implies a'_B(v) = 0$ *for* v *in this interval. Thus,* $a_B(x_i) = a_B(x_{i+1})$*.*

Remark 8. *For all* $i \in \{1, 2, \cdots, M\}$*, we have* $u_A(x_i) = u_B(x_i)$*.*

We now prove a lemma that forms the backbone of our inductive argument:

Lemma 2. *For all* $i \in \{1, 2, \cdots, M-1\}$*, we have* $a_A(x_i) > a_B(x_i) \iff a_A(x_{i+1}) < a_B(x_{i+1})$*. Similarly, we have* $a_A(x_i) < a_B(x_i) \iff a_A(x_{i+1}) > a_B(x_{i+1})$

Proof. Note that either $a_A(x_i) = a_A(x_{i+1})$ or $a_B(x_i) = a_B(x_{i+1})$ by Remark 6 and Remark 7. We only prove $a_A(x_i) > a_B(x_i) \iff a_A(x_{i+1}) < a_B(x_{i+1})$ for the case $a_A(x_i) = a_A(x_{i+1})$ and omit the other (symmetric) cases. Since $a_A(x_i) = a_A(x_{i+1})$, we have

$$u_A(x_i) = u_A(x_{i+1}) + a_A(x_i) \cdot (x_i - x_{i+1}) = a_A(x_{i+1}) \cdot (\underline{x_i}_B - x_{i+1}) + a_A(x_i) \cdot (x_i - \underline{x_i}_B).$$

We also have, by the structure of the ironed intervals for the item B,

$$u_B(x_i) = u_B(x_{i+1}) + a_B(x_{i+1}) \cdot (\underline{x_i}_B - x_{i+1}) + a_B(x_i) \cdot (x_i - \underline{x_i}_B)$$

Now, since the utilities at all points x_i is the same for both items A and B (Remark 8), the fact that $a_A(x_i) > a_B(x_i)$ is equivalent to $a_A(x_i) \cdot (x_i - \underline{x_i}_B) > a_B(x_i) \cdot (x_i - \underline{x_i}_B)$ which is equivalent to $a_A(x_{i+1}) \cdot (\underline{x_i}_B - x_{i+1}) < a_B(x_{i+1}) \cdot (\underline{x_i}_B - x_{i+1})$ which, in turn, is equivalent to $a_A(x_{i+1}) < a_B(x_{i+1})$. ☐

Finally, we prove Theorem 3.

Proof of Theorem 3. At x_M, we have that

$$u_A(x_M) = a_A(x_M) \cdot (x_M - \underline{x_M}_A) \qquad \text{and} \qquad u_B(x_M) = a_B(x_M) \cdot (x_M - \underline{x_M}_B).$$

Since $\underline{x_M}_B > \underline{x_M}_A$ and $a_A(x_M) > 0$, then to ensure that $u_A(x_M) = u_B(x_M)$ (Remark 8), we must have $a_B(x_M) > a_A(x_M)$. However, with this fact, Lemma 2 says that $a_B(x_i) > a_A(x_i)$ and $a_B(x_{i+1}) > a_A(x_{i+1})$ in alternation.

Since $a_A(\cdot)$ and $a_B(\cdot)$ are non-decreasing sequences, they can only alternate if they have $\Omega(M)$ distinct elements.

By Theorem 4, there exists a feasible primal that satisfies complementary slackness. The primal algorithm constructs a mechanism with menu complexity at least M and satisfies complementary slackness, hence this dual is in fact optimal. □

Corollary 9. *This idea gives a lower bound for Multi-Unit Pricing as well.*

We expand on this on Section 4.12.

4.8 Menu Complexity is Finite: Characterizing the Optimal Mechanism via Duality

In this section, we'll characterize the optimal mechanism for three items $\{A, B, C\}$ with structure $A \succ C$, $B \succ C$, and $A \not\succ B, B \not\succ A$. While our approach will be algorithmic, our focus isn't to actually run this algorithm or analyze its runtime. We'll merely use the algorithms to deduce structure of the optimal mechanism. We prove essentially that the interleaving of ironed intervals used in the construction of the previous section is the worst case (in terms of menu complexity of the optimal mechanism). Still, in order to possibly prove this, we need to at minimum find an optimal mechanism for every possible instance.

Our approach is the following: we propose a *primal recovery algorithm* that, given a dual (λ, α), produces a primal solution that (1) satisfies complementary slackness with the dual and (2) has finite menu complexity. Obviously, the algorithm can't possibly succeed for every input dual (as some duals are simply not optimal for any instance). But we show that whenever the algorithm fails, the dual has some strange structure (elaborated below). We then show that the best dual (which is optimal and always exists, definition below) never admits these strange structures, and therefore the algorithm always succeeds when given the best dual as input.

Definition 10 (Best Dual)**.** We define the best dual of an instance with three partially-ordered items to be the (λ, α) satisfying the following:

1. First, (λ, α) is optimal: $(\lambda, \alpha) \in \arg\min\{\sum_{G\in\{A,B,C\}} \int_0^H f_G(v)\cdot\max\{0, \Phi_G^{\lambda,\alpha}(v)\}dv\}$.

2. Among (λ, α) satisfying (1), (λ, α) has the *fewest ironed intervals* of virtual value zero. That is, (λ, α) minimizes $|\mathcal{I}(\lambda, \alpha)| = |\{\underline{x}_G \mid (x, G) \in [0, H] \times \{A, B, C\}, \Phi_G^{\lambda,\alpha}(v) = 0\}|$.

3. Among (λ, α) satisfying (2), (λ, α) has the *lowest positives* (lexicographically ordered). That is, (λ, α) minimizes $\bar{r}_A$, followed by $\bar{r}_B$, followed by $\bar{r}_C$.

Definition 11. A *double swap* exists when there are consecutive points (x, A) and (y, B) in a chain, and there is flow into A for $v \in [\underline{x}_A, y)$. See Figure 4.14.

Definition 12. An *upper swap* occurs when there is flow into (x, A) and (y, B) where $x > \bar{r}_A > y > \bar{r}_B$. See Figure 4.15.

Proposition 10. *The best dual has no double swaps or upper swaps.*

The full proof of Proposition 10 appears below. The high-level approach is that whenever a double swap or upper swap exists, we can exploit this structure to modify the dual variables. This creates a better dual solution (with respect to definition 10) and proves that (2) or (3) respectively must not have held for the original dual.

Theorem 4. *For any best dual solution, we can find a primal with finite menu complexity that satisfies complementary slackness (and is therefore optimal).*

A full proof appears below, but the high-level approach is explained in the following.

Proof Sketch of Theorem 4. (No bad structures exist in best duals.) First, we try to satisfy the necessary complementary slackness system of equations as in Section 4.7, and identify all possible barriers to solutions existing. These barriers are exactly double swaps or upper swaps, which are not found in best duals by Proposition 10.

(Inductive primal recovery algorithm.) Without these barriers, an inductive argument shows that we can indeed find an allocation rule that satisfies all of the complementary slackness conditions. Every dual has a (possibly empty) top chain, and each point in the chain has another set of preferability constraints for that item, along with the constraint that the allocation is constant. We use induction to handle one point in the chain at a time. (See Figure 4.18 in Section 4.8.) We take the partially-constructed allocation that satisfies the constraints for the chain so far, scale it down (and thus continue to satisfy the constraints), and then solve for the allocation probability that will satisfy the new constraints given by this point in the chain. As shown in Section 4.7, this requires choosing a different allocation probability at the bottom of each ironed interval in the chain, but we show that this is sufficient, giving menu complexity at most the length of the chain $+1$.

(Finite menu complexity.) The other interesting part not addressed in Section 4.7 is what to do if there is a chain of countably infinite length (which can certainly exist). Naively following our algorithm would indeed result in a primal of countably infinite menu complexity. But, because the sequence of chain points is monotonically decreasing (and lower bounded by zero), they must converge to some value v. If they converge, and the chain is indeed infinitely long, then neither A nor B can possibly be ironed at v, and we can simply set price v for both items instead. □

We begin below by reviewing properties of the dual previously observed in Fiat et al. [2016]; Devanur and Weinberg [2017]. Throughout this section we'll reference the "best" dual. While multiple optimal duals might exist, we'll be interested in a specific tie-breaking among them (and refer to the one that satisfies these conditions as "best").

Theorem 11 (Devanur and Weinberg [2017]). *The best dual* (λ, α) *satisfies the following:*

- *(Proper monotonicity)* $(f_G \cdot \Phi_G^{\lambda,\alpha})(\cdot)$ *is monotone non-decreasing, for all* v.
- *(No-boosting)* $\Phi_G^{\lambda,\alpha}(v) \geq 0$ *for all* G *such that there exists a* $G' \succ G$.
- *(No-rerouting)* $\Phi_G^{\lambda,\alpha}(v) > 0 \Rightarrow \alpha_{G,G'}(v) = 0$ *for all* G'.
- *(No-splitting)* $\lambda_G(v) > 0 \Rightarrow \alpha_{G,G'}(v) = 0$ *for all* G'.

Returning to our three-item example, prior work nicely characterizes the flow coming out of C in the optimal dual: No-boosting tells us that we must always send flow out of (v, C) into somewhere whenever $\Phi_C^{\lambda,\alpha}(v) < 0$ (in order to bring it up to 0). No-rerouting tells us that we can never send flow out of (v, C) if $\Phi_C^{\lambda,\alpha}(v) > 0$. No-splitting tells us that we never send flow out of the middle of an ironed interval. But, we still need to decide whether to send this flow into A or B. This is the novel part of our analysis.

Proof of Proposition 10. By Definition 10, we know that a best dual has the minimum number of ironed intervals amongst all optimal duals. Similarly, a best dual has the lowest positives amongst all optimal duals. We prove the proposition using two lemmas. The first lemma proves that a best dual can't have double swaps:

Lemma 3. *The optimal dual that has the minimal number of ironed intervals does not contain any double swaps.*

First, we discuss why this structure would cause a problem for how we're used to satisfying complementary slackness conditions. Complementary slackness forces that in the ironed intervals $[\underline{z}_B, \overline{z}_B]$ and $[\underline{y}_A, \overline{y}_A]$, the allocation is constant, and thus utility in these regions is linear. However, no linear utility functions can satisfy the preferability constraints of having utility that is higher for item A, then B, then A, as illustrated on the left in Figure 4.14.

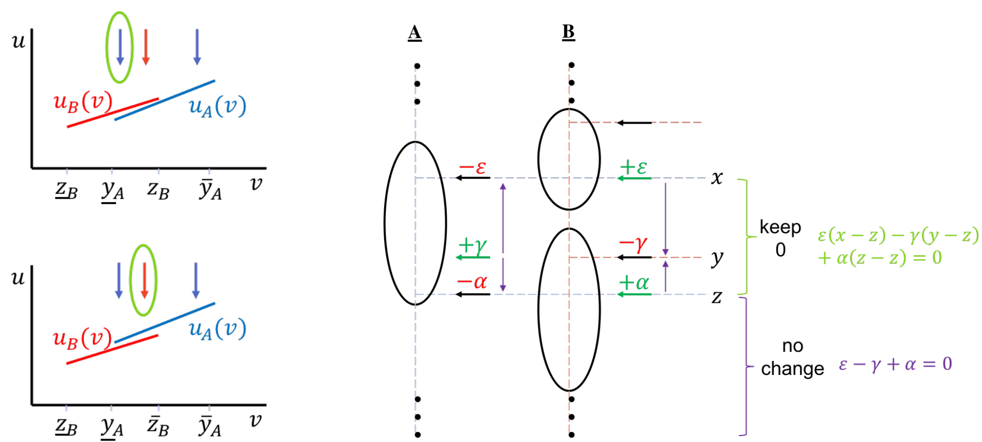


Figure 4.14: Left: Complementary slackness forces linear utility in ironed intervals. For any choice of linear utility functions, we cannot satisfy the preferability constraints imposed by the double swap for item A, then B, then A in this region. The violated constraint corresponds to the circled arrow. Right: The operation used in the proof of Lemma 3, using a double swap to maintain virtual welfare and create fewer ironed intervals.

Proof. Proof by contradiction. Suppose that somewhere in the top chain, some point in the chain (x, A) is succeeded by (y, B) and $\alpha_{C,A}(z) > 0$ for some $z \in (\underline{x}_A, y)$, creating a double swap. We consider the following operation (depicted on the right in Figure 4.14) that pushes flow down within the ironed interval $[\underline{x}_A, \overline{x}_A]$ and does the reverse on B, yet negates the change in flow at z to maintain the virtual values below here. Move ε flow from (x, A) to (x, B). Move γ flow from (y, B) to (y, A). Move α flow from (z, A) to (z, B). We will set

$$\alpha = \left(\frac{x-y}{y-z}\right)\varepsilon \qquad \text{and} \qquad \gamma = \left(1 + \frac{x-y}{y-z}\right)\varepsilon.$$

First, this ensures that $\varepsilon - \gamma + \alpha = 0$, and thus for $v \leq z$, $\hat{\Phi}_A^{\lambda,\alpha}(v) = \Phi_A^{\lambda,\alpha}(v)$ as well as $\hat{\Phi}_B^{\lambda,\alpha}(v) = \Phi_B^{\lambda,\alpha}(v)$. Second, this ensures that $\varepsilon(x-z) - \gamma(y-z) = 0$, keeping the average virtual value from z to x the same for both items.

$$\begin{aligned}
\int_z^x f_A(v)\hat{\Phi}_A^{\lambda,\alpha}(v)dv &= \int_z^y f_A(v)(\Phi_A^{\lambda,\alpha}(v) + \varepsilon - \gamma)dv + \int_y^x f_A(v)(\Phi_A^{\lambda,\alpha}(v) - \gamma)dv \\
&= \int_z^x f_A(v)\Phi_A^{\lambda,\alpha}(v)dv + \varepsilon(x-z) - \gamma(y-z) \\
&= \int_z^x f_A(v)\Phi_A^{\lambda,\alpha}(v)dv
\end{aligned}$$

However, the virtual values in $[y, x]$ are increasing for item A and decreasing for item B, and likewise those in $[z, x)$ are decreasing for item A and increasing for item B. If we choose ε small enough as to not uniron the interval $[\underline{x}_A, \bar{x}_A]$, the change gets spread around the interval and the interval remains all zeroes. However, for item B, the interval $[\underline{y}_B, \bar{y}_B]$ becomes positive while the region above becomes negative. Since the average of both regions is the same and there is now a non-monotonicity, the regions will be ironed together, creating a larger ironed interval with virtual value zero.

Since the virtual welfare of the dual hasn't changed, but we have reduced the number of ironed intervals, then we did not start with an optimal dual with the fewest possible ironed intervals, deriving a contradiction. □

The second lemma proves that a best dual can't have upper swaps:

Lemma 4. *The optimal dual that has the lowest positives does not contain any upper swaps.*

Proof. Proof by contradiction. Suppose an upper swap exists. Then (as depicted on the right in Figure 4.15) we can push up α flow from (y, B) to (x, B), causing $f_B(v)\hat{\Phi}_B^{\lambda,\alpha}(v) = f_B(v)\Phi_B^{\lambda,\alpha}(v) - \alpha$ for $v \in [y, x]$ and improving virtual welfare by

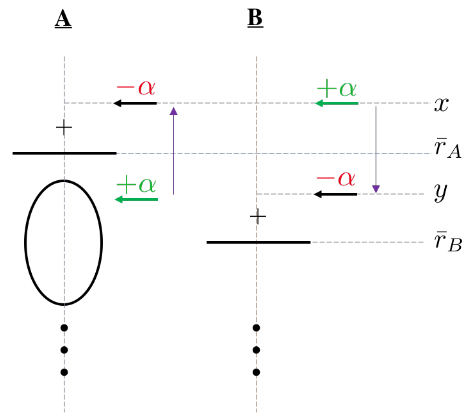


Figure 4.15: The operation used in the proof of Lemma 4, using a upper swap to maintain virtual welfare and create lower positives.

$\alpha(x-y)$. To leave the flow out of item C unchanged, we balance this out by pushing α flow down from (x, A) to (y, A), causing $f_A(v)\hat{\Phi}_A^{\lambda,\alpha}(v) = f_A(v)\Phi_A^{\lambda,\alpha}(v) + \alpha$ for $v \in [y, x]$.

If y is unironed at A, that is, $\bar{y}_A = y$, or if $\bar{y}_A < \bar{r}_A$, then by choosing $\alpha = -f_A(\bar{y}_A)\Phi_A^{\lambda,\alpha}(\bar{y}_A)$, this will cause $\hat{\bar{r}}_A = \bar{y}_A$, lowering the positives.

Alternatively, if y is ironed up to $\bar{r}_A$ such that $\bar{y}_A = \bar{r}_A$, then we can choose a very small α to keep the interval $[\underline{y}_A, \bar{r}_A]$ ironed, making the whole interval positive and causing $\hat{\bar{r}}_A = \underline{y}_A$, lowering the positives. The dual will only increase by $\alpha(x-y)$, even when the values are ironed around, as ironing preserves virtual welfare. This is canceled out by the improvement in virtual welfare from item B. Then we have maintained virtual welfare but lowered the positives, showing that this dual solution could not have had the lowest positives. □

Lemma 3 and Lemma 4 comprise the proof of Proposition 10. □

Now we prove that our primal recovery algorithm always succeeds in finding an optimal primal (that satisfies complementary slackness) when given a best dual.

Proof of Theorem 4. First, consider the case where there exists some point v where $\Phi_A^{\lambda,\alpha}(v) = \Phi_B^{\lambda,\alpha}(v) = 0$, and v is unironed both in A and in B. Then we simply set v as a price for both A and B, automatically satisfying the complementary slackness conditions of flow into A or B, as both are equally preferable. Since both items A

and B, have the same allocation rule, the instance degenerates into a FedEx instance. Thus, an optimal allocation rule for the item C can be determined.

Otherwise, the dual solution contains no point v as described in the first case, meaning that ironed intervals interleave throughout the region of zero virtual values. This means that, if without loss of generality $\bar{r}_A > \bar{r}_B$, that $\bar{r}_B = x$ must sit in an ironed interval $[\underline{x}_A, \bar{x}_A]$ on A.

If the top chain is empty, then we have $\bar{r}_A > \bar{r}_B > \underline{x}_A$ with no flow into A for any $v \in [\underline{x}_A, \bar{x}_A]$. Then, setting

$$a_A(v) = \begin{cases} 1 & v \geq \bar{x}_A \\ \frac{\bar{r}_A - \bar{r}_B}{\bar{r}_A - \underline{x}_A} & v \in [\underline{x}_A, \bar{x}_A) \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad a_B(v) = \begin{cases} 1 & v \geq \bar{r}_B \\ 0 & \text{otherwise} \end{cases}$$

makes both options equally preferable for all v except for $v \in [\underline{x}_A, \bar{x}_A]$, where reporting B is strictly preferable, but this does not violate complementary slackness by the assumption that the top chain is empty.

Otherwise, the top chain is non-empty. A dual gives a system of utility inequalities via complementary slackness which the allocation rule must satisfy. Instead, we can solve a system of utility equalities given by the chain via induction on the length of the top chain, and this will imply a solution that satisfies all of the inequalities. More specifically, the following will hold for top chains of all lengths:

1. The allocation rule will only increase at the bottom of ironed intervals in the chain. That is, if the allocation rule increases at z, so $a'_A(z) > 0$, then z must be the bottom of an ironed interval for a point (x, A) in the top chain, thus $z = \underline{x}_A$, and $a_A(x) = a_A(\underline{x}_A)$.

2. We will fully allocate to all positive virtual values. That is, $a_A(\bar{r}_A) = a_B(\bar{r}_B) = 1$.

3. If (x, A) is followed by (y, B) in the chain, then $a_A(x) = a_A(\underline{x}_A) > a_B(y) = a_B(\underline{y}_B)$.

4. At any point (x, A) in the top chain, we will have $u_A(x) = u_B(x)$.

5. An alternative solution can, for the first point in the chain (x, A), vary $a_A(\underline{x}_A)$ such that the utility constraint is a strict inequality $u_A(x) > u_B(x)$, and instead

we have equality at $\bar{r}_A$: $u_A(\bar{r}_A) = u_B(\bar{r}_A)$. This gives an equal expected price for the two items, and equal utility for all values $v \geq \bar{r}_A$.

To satisfy complementary slackness, for any type (x, A) with flow in, it must be that $u_A(x) \geq u_B(x)$. We now show why (3-4) imply that complementary slackness will be satisfied everywhere.

Consider a subsequence of points in the chain: $(x, B), (y, A), (z, B)$, hence $y > \underline{x}_B$ and $z > \underline{y}_A$. Then $a_B(x) > a_A(y) > a_B(z)$ by (3). Since $u_A = u_B$ for every point in the chain and a larger allocation rule implies a larger change in utility, we can deduce that $u_A(v) \geq u_B(v)$ for all $v \in [z, y]$.

- For $v \in (\underline{y}_A, \underline{x}_B)$, we have that $a_A(v) > a_B(v)$, and since $u_A(z) = u_B(z)$, then $u_A(v) \geq u_B(v)$ in this region.
- For $v \in (\underline{x}_B, y)$, we have that $a_B(v) > a_A(v)$, and since $u_A(y) = u_B(y)$, then $u_A(v) \geq u_B(v)$ in this region.
- By definition of a double swap, there is no $v \in [\underline{y}_A, z)$ such that there is flow into (v, A). Likewise, there is no $v \in [\underline{x}_B, y)$ such that there is flow into (v, B).

Hence all possible complementary slackness conditions are satisfied.

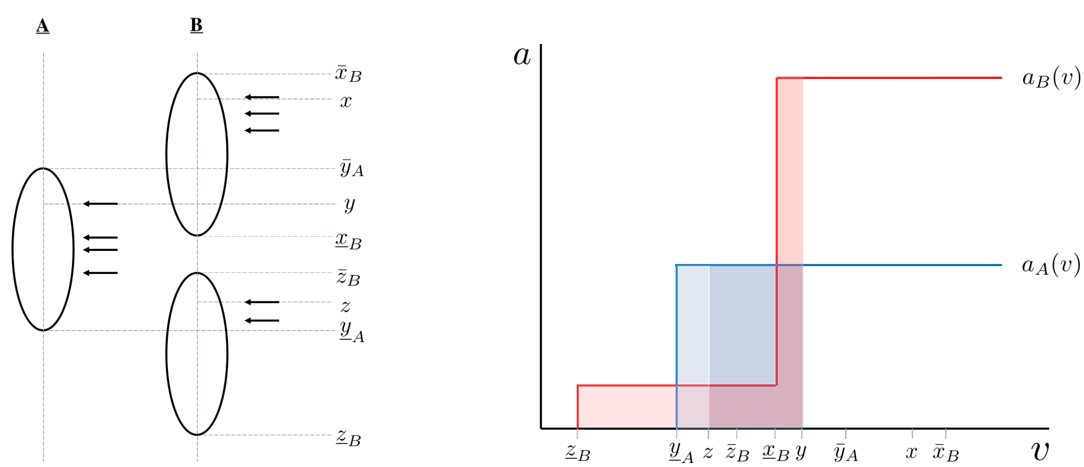


Figure 4.16: Left: A candidate dual (with no double or upper swaps); part of a chain. Right: An allocation that satisfies complementary slackness up to value y, satisfying equal preferability at z and y and preferability at all points with flow in.

We now show that these sufficient properties hold by induction. As a base case, consider when there is one point in the top chain, which without loss is (x, A). By definition of the top chain, $\bar{r}_A > x > \bar{r}_B > \underline{x}_A$ and there is flow into item A at x, which is in ironed interval $[\underline{x}_A, \bar{x}_A]$. We can set $a_A(\underline{x}_A) = \frac{x-\bar{r}_B}{x-\underline{x}_A}$ and set $a_A(\bar{r}_A) = a_B(\bar{r}_B) = 1$. Then

$$u_A(x) = a_A(\underline{x}_A) \cdot (x - \underline{x}_A) = \frac{x - \bar{r}_B}{x - \underline{x}_A} \cdot (x - \underline{x}_A) = 1(x - \bar{r}_B) = u_B(x).$$

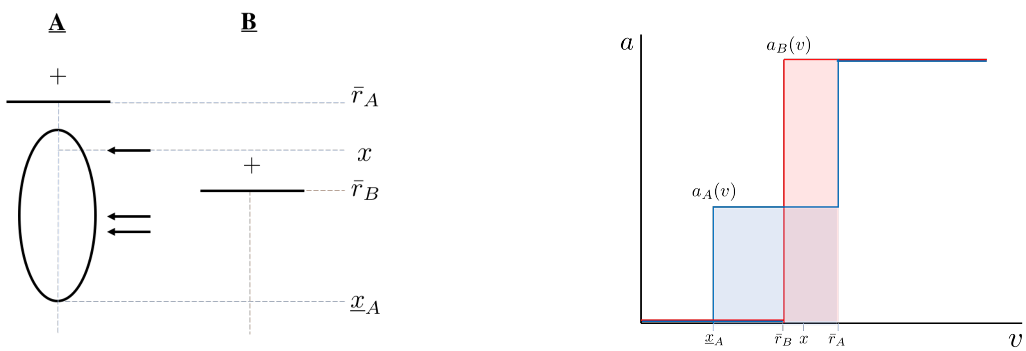


Figure 4.17: Left: The base case of a candidate dual with an empty chain. Right: An allocation that satisfies complementary slackness.

Then conditions (1-4) are met. To satisfy (5), we can instead set $a_A(\underline{x}_A) = \frac{\bar{r}_A-\bar{r}_B}{\bar{r}_A-\underline{x}_A}$. Then

$$u_A(\bar{r}_A) = a_A(\underline{x}_A) \cdot (\bar{r}_A - \underline{x}_A) = \frac{\bar{r}_A - \bar{r}_B}{\bar{r}_A - \underline{x}_A} \cdot (\bar{r}_A - \underline{x}_A) = 1(\bar{r}_A - \bar{r}_B) = u_B(\bar{r}_A).$$

For the inductive hypothesis, suppose for any chain of length n, we have allocation rules such that (1-5) hold.

Now consider a chain of length $n + 1$. Without loss of generality, let (x, A) be the top point in the chain, where x sits in the ironed interval $[\underline{x}_A, \bar{x}_A]$, and this point is proceeded by (y, B) which sits in $[\underline{y}_B, \bar{y}_B]$, hence $\bar{r}_A > x > \bar{r}_B$ and $y > \underline{x}_A$ by definition of the chain.

By the inductive hypothesis, we can come up with allocation rules $a_A(\cdot)$ and $a_B(\cdot)$ that satisfy complementary slackness to the same chain without the highest point (x, A), and will have $a_A(\underline{x}_A) = a_B(\bar{r}_B) = 1$. We construct an allocation rule

$\hat{a}$ for the top chain of size $n+1$ as follows; this is depicted in Figure 4.18. Let $\lambda = \frac{x-\bar{r}_B}{x-y-a_B(\underline{y}_B)(\bar{r}_B-y)} < 1$. Then let

$$\hat{a}_A(v) = \begin{cases} 1 & v \geq \bar{r}_A \\ \lambda a_A(v) & \text{otherwise} \end{cases} \qquad \text{and} \qquad \hat{a}_B(v) = \begin{cases} 1 & v \geq \bar{r}_B \\ \lambda a_B(v) & \text{otherwise.} \end{cases}$$

This clearly satisfies (1-3). To show that (4) holds, we observe that at any previous point of concern $v < \bar{r}_B$, we had $u_A(v) = u_B(v)$. Now at those points, we have $\hat{u}_A(v) = \int_0^v \hat{a}_A(v)dv = \lambda \int_0^v a_A(v)dv = \lambda u_A(v)$. This holds for $\hat{u}_B(v) = \lambda u_B(v)$ as well. Thus, complementary slackness is still satisfied at all previous points $v \leq \bar{r}_B$; we only need to check equal utility at x.

$$\hat{u}_A(x) = \hat{u}_A(y) + \hat{a}_A(\underline{x}_A)(x-y) = \lambda u_A(y) + \lambda \cdot 1 \cdot (x-y)$$
$$\hat{u}_B(x) = \hat{u}_B(y) + \hat{a}_B(\underline{y}_B)(\bar{r}_B - y) + \hat{a}_B(\bar{r}_B)(x-\bar{r}_B) = \lambda u_B(y) + \lambda \cdot a_B(\underline{y}_B)(\bar{r}_B - y) + 1 \cdot (x - \bar{r}_B)$$

Then to have $\hat{u}_A(x) = \hat{u}_B(x)$, since $u_A(y) = u_B(y)$, we require that

$$\lambda(x-y) = \lambda \cdot a_B(\underline{y}_B)(\bar{r}_B - y) + 1 \cdot (x - \bar{r}_B).$$

The solution here is exactly the λ defined above.

Alternatively, by replacing x with $\bar{r}_A$, thus setting $\lambda = \frac{\bar{r}_A-\bar{r}_B}{\bar{r}_A-y-a_B(\underline{y}_B)(\bar{r}_B-y)}$, we get a solution that has $u_A(x) > u_B(x)$ and $u_A(\bar{r}_A) = u_B(\bar{r}_A)$ as required in (5).

Thus we have ensured that for top chains of all lengths, we can give an allocation rule that satisfies complementary slackness for all values from the bottom to the top of the chain. For v below the chain, $u_B(v) = u_A(v) = 0$, so we automatically satisfy complementary slackness. Above the chain, if we have used the alternate solution that (5) guarantees exists, we automatically satisfy complementary slackness for $v \geq \bar{r}_A$. This would only fail if there is flow into item B for $v \in [x, \bar{r}_A)$—that is, if the dual contains a upper swap, but by assumption it does not. Then for any dual solution with no double swaps or upper swaps, this algorithm gives an allocation rule that satisfies complementary slackness.

We prove that the menu complexity of the mechanism output by this algorithm is finite below:

Claim 6. The menu complexity is always finite.

Proof. Proof by contradiction. Suppose that there exists an instance such that the mechanism output by the algorithm has infinite menu complexity.

Note that this can only happen if the length of the top chain is infinity. Thus, there exists a sequence of points $(x_1, A), (x_2, B), (x_3, A), \cdots$ such that the point (x_i, A) is inside an ironed interval $[\underline{x_i}_A, \overline{x_i}_A]$ and $x_{i+1} \geq \underline{x_i}_A$. Analogous claims hold for an element (x_{i+1}, B) in the chain.

Thus, we have

$$x_1 \geq \overline{r}_B \geq \overline{x_2}_B \geq x_2 \geq \underline{x_1}_A \geq \overline{x_3}_A \geq x_3 \geq \cdots$$

Since the infinite sequence $x_1, x_2, \cdots$ is monotone and bounded, it converges to a limit, say x^*. Observe that x^* satisfies $\Phi_A^{\lambda,\alpha}(x^*) = \Phi_B^{\lambda,\alpha}(x^*) = 0$ and is unironed. This is because points arbitrarily close to it are unironed and are zeroes of $\Phi_A^{\lambda,\alpha}(\cdot)$ and $\Phi_B^{\lambda,\alpha}(\cdot)$. However, in this case, our algorithm just sets the price x^* and thus has constant menu complexity, a contradiction.

□

□

4.9 An Exact Characterization Under the Assumption of DMR

Recall from Subsection 4.3.2 that when the distributions are DMR, $\lambda_G(v) = 0$ for all (v, G). Our main result is the following:

Theorem 12. *Consider any partially-ordered preferences for items. If the marginal distribution for each item is DMR, the optimal mechanism is deterministic.*

4.9.1 Intuition

For a partial ordering given by $\mathcal{G}, \succ$ represented as a DAG such that the marginal distributions satisfy DMR, the optimal mechanism will set prices as follows. The algorithm is analogous to that in FedEx.

- For items G that are leaf nodes in the DAG, set $p_G = r_G$.

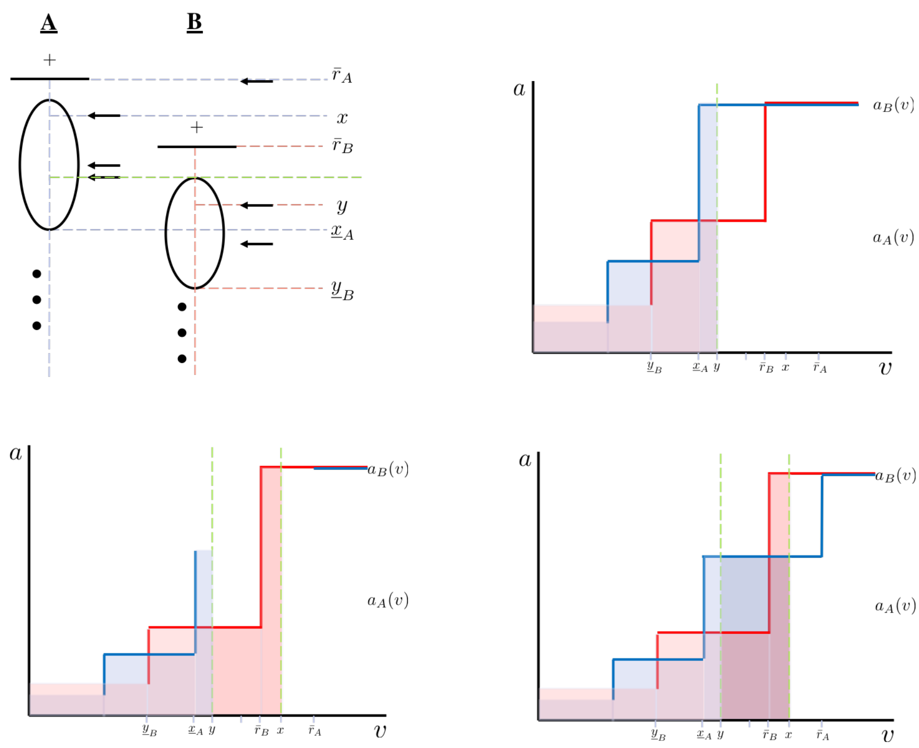


Figure 4.18: Top Left: A top chain from a candidate dual. We use the inductive hypothesis on the chain of one size smaller (below the green line). Top Right: The allocation rule from the inductive hypothesis that satisfies all CS constraints on the smaller chain (below the green line). Bottom Left: The scaled allocation rule, requiring preferability of A between the green lines. Bottom Right: The allocation rule that satisfies these preferability constraints.

- From better items to worse items, in reverse depth from the leaf nodes, we will define a least upper bound on each node's price based on the prices set for the better items. We define $\bar{p}_G = \min_{G' \in N^+(G)} \bar{p}_{G'}$ to be the least upper bound on G's price. Then set a price of $p_G = \min\{\bar{p}_G, r_G\}$ for G.

In our pricing algorithm, nodes G are limited by the smallest r_A for any A that they have a directed path to. From complementary slackness, every r_A that a node G has a path to is an upper bound on the price that can be set for G, so the smallest of these upper bounds is the most limiting. We define $\bar{p}_G$ to be this smallest upper bound, and we define L_G to be the nodes from G's 1-out-neighborhood who are also constrained by this upper bound. Thus, if we follow the sets L_G, we result in all of the limiting nodes with $r_A = \bar{p}_G$.

When we send flow out of G, we aim to send it along the path to the node that limits G's price the most. We do this recursively, sending from G to the most limiting neighbor, and from there to that node's most limiting neighbor, splitting the flow equally if there are several limiting neighbors. This raises (does not lower) the limiting reserve. We update regularly to ensure that we are always sending flow to the now-limiting reserve, raising it, and thus minimizing the constraints on G. This is almost exactly the construction: the only caveat is that we should never send flow out of an item B at v where $f_B(v)\Phi_B^{\lambda,\alpha}(v) > 0$. If we send into a B along the path where this is the case, we instead send out at $r_B < v$.

Formally, we set the dual variables according to the following algorithm:

Dual variable construction:
Base case: For leaf nodes, there is nowhere to send flow. $\bar{p}_A = r_A$.
for all nodes A starting from the leaf nodes and in increasing reverse depth (# edges from leaf nodes) **do**
  $\bar{p}_A = \min_{B \in N^+(A)} \bar{p}_B$
  For all v from r_A down to 0, determine the minimal amount of flow out σ_A such that $\varphi_A(v) = 0$.
  for v from 0 to r_A **do**
    Update$(A, v, \sigma_A(v))$
  end for
end for

Update(A, v, γ):
Let $L_A := \{\text{argmin}_{B \in N^+(A)} \bar{p}_B\}$.
for all $B \in L_A$ **do**
  Send $\alpha_{A,B}(v) = \frac{1}{|L_A|}\gamma$.
  Update$(B, \min\{v, r_B\}, \gamma)$.
end for

The key idea is that the price of G is limited by the smallest r_A where A is some better item than G, that is, there is a path from G to A in the DAG representing the partial ordering. As we send flow along the path to A, we raise r_A and it becomes less limiting. We will maintain a set of the items that limit G the most, S_G, which are precisely the items A such that $r_A = \bar{p}_G$. Because we are in the continuous setting, sending flow is a continuous process, so the most limiting item never discretely jumps up higher and becomes no longer limiting. Instead, all limiting items stay in the set S_G and this set grows as the upper bounds raise up and become less limiting.

We define $L_G \subseteq N^+(G)$ to be the items such that, for all $B \in L_G$, there exists v such that $\alpha_{G,B}(v) > 0$. What this means is that $\bar{p}_G = \bar{p}_B$, and B is on the path (if not the end of the path) from G to a limiting item $A \in S_G$. We will track the updated $\bar{p}_G$ with r. If $A \in S_G$, then $f_A(r)\Phi_A^{\lambda,\alpha}(r) = 0$, and if B is on a path to some limiting A, then $f_B(r)\Phi_B^{\lambda,\alpha}(r) \leq 0$.

In every step, we decrease the amount of flow to send and increase r. We terminate when there is no flow left to send. The point r only increases and the set S_G only increases. We maintain the above properties.

First, we set the flow out of G:

$$\sum_{A \in N^+(G)} \alpha_{G,A}(v) = \begin{cases} 0 & v > \bar{r}_G \\ -\hat{R}_G''(v)/f_G(v) & v \leq \bar{r}_G. \end{cases}$$

Lemma 5. *For every G, we can always send σ_G out of G distributed among $N^+(G)$ such that*

1. *If $\alpha_{G,B}(x) > 0$ for any x then $B \in L_G$.*
2. *If $B \in L_G$ then $f_B(r)\Phi_B^{\lambda,\alpha}(r) = 0$.*
3. *If $B \in N^+(G) \setminus L_G$ then $r < \overline{r}_B$ and thus $f_B(r)\Phi_B^{\lambda,\alpha}(r) \leq 0$.*

Proof. Suppose we have $\sigma_G(v)$ flow to send at v. Let $Z = \operatorname{argmin}_{B \in N^+(G) \setminus L_G} \overline{p}_B$ be the next possible upper bound to hit.

We choose ε such that by sending σ_G flow along paths to all items in S_G with correct proportions, we will maintain S_G and raise r by ε. That is,

$$\sum_{A \in S_G} f_A(r+\varepsilon)\Phi_A^{\lambda,\alpha}(r+\varepsilon) = \sigma.$$

If $r + \varepsilon < \overline{p}_Z$, we can send this flow without growing S_G. Let $\mathcal{P}(G, A)$ denote the edges forming every path from G to A. For every $(C, D) \in \mathcal{P}(G, A)$ for some $A \in S_G$, we set

$$\alpha_{C,D}(v) = \sum_{A \in S_G : (C,D) \in \mathcal{P}(G,A)} f_A(r+\varepsilon)\Phi_A^{\lambda,\alpha}(r+\varepsilon) \quad \forall A \in L_G$$

we ensure that after this update, $f_A(r+\varepsilon)\Phi_A^{\lambda,\alpha}(r+\varepsilon) = 0$ for all $A \in S_G$. Update $r \leftarrow r+\varepsilon$. Clearly (2) holds, and (3) holds since $r < \overline{p}_Z < \overline{p}_B$ for all $B \in N^+(G) \setminus L_G$.

Otherwise, $r + \varepsilon \geq \overline{p}_Z$ and $v \geq \overline{p}_Z$. Then we instead choose $\varepsilon = \overline{p}_Z - r$ and make the same update described above, then add Z to L_G and add the item Y that is limiting Z, that is, Y such that $\overline{p}_Z = R_Y$, to S_G. Note that we have sent positive flow, but the flow sent is $< \sigma$. After the update, we will have $r \leftarrow r + \varepsilon = \overline{p}_Z$ and $f_A(r)\Phi_A^{\lambda,\alpha}(r) = 0$ for all $A \in S_G$, including Y. Then clearly (2) holds, and (3) holds since $r = \overline{p}_Z < \overline{p}_B$ for all $B \in N^+(G) \setminus L_G$.

Finally, (1) holds in both cases as we only send flow to elements of S_G and S_G is non-decreasing. □

Lemma 6. *For every v and G, our choice of $\alpha_{G,A}(w)$ for all $w \in [0, H]$, $A \in N^+(G)$ maintains $\lambda_G(v) = 0$ for all v.*

Proof. Since the flow out of G is chosen exactly to bring all virtual values to 0 below $\bar{r}_G$, no non-monotonicities are caused. □

Lemma 7. *For every v and G, any choice of $\alpha_{A,G}(w)$ for all $w \in [0, H]$, $A \in N^-(G)$ maintains $\lambda_G(v) = 0$ for all v.*

Proof. Suppose we get flow α into G at x. Every value $v \leq x$ has $f_G(v)\Phi_G^{\lambda,\alpha}(v)$ decrease by α while this remains unchanged for $v > x$, causing no non-monotonicities. □

Lemma 8. *The following deterministic allocation rule always satisfies complementary slackness with the dual: Set $p_G = \min\{r_G, r_A : A \in S_G\}$.*

Proof. From DMR and our setting of λ, we will have $\lambda_G(v) = 0$ for all (v, G), automatically satisfying complementary slackness for these variables. Further, even after sending α flow, $f_G(\cdot)\Phi_G^{\lambda,\alpha}(\cdot)$ will be properly monotone for all G by Lemma 6 and Lemma 7.

First, we verify that the when we set a price, the virtual values are 0 at that price, so we have the freedom to do so.

By Lemma 5, $f_A(r)\Phi_A^{\lambda,\alpha}(r) = 0$ for all $A \in S_G$. Of course, by definition of $\bar{r}$, $f_A(\bar{r}_A)\Phi_A^{\lambda,\alpha}(\bar{r}_A) = 0$. In addition, by definition of the flow out of G, $f_G(v)\Phi_G^{\lambda,\alpha}(v) = 0$ for all $v \leq \bar{r}_G$ so $f_G(r)\Phi_G^{\lambda,\alpha}(r) = 0$. Then all of the prices posted are viable.

It remains to choose a mechanism that satisfies complementary slackness with the α variables. If $\alpha_{G,B}(v) > 0$ for some v then we know that (1) $B \in L_G$ and (2) $v < \bar{r}_G$.

By Lemma 5, the variable $\alpha_{G,B}(v) > 0$ for any v if and only if $v \in L_G$, a monotone increasing set as v increases. In this case, then $\bar{p}_B = \bar{p}_G$ and both are set at this price, satisfying $u_G(v) = u_B(v)$ for all v and automatically satisfying complementary slackness. □

4.10 Equivalence with Single-Minded Valuations

In the introduction, we note the following observation.

Observation 13. *The partially-ordered setting is* equivalent *to the single-minded setting.*

First, we define the single-minded setting.

Definition 13. In a *single-minded setting*, a seller determines how to sell any bundle of m items. A buyer has a (value, bundle) pair (v, B) where $B \in 2^{[m]}$ is any subset of items. The pair (v, B) is drawn from a joint probability distribution over $[0, H] \times 2^{[m]}$ where H is the maximum possible value of any bidder for any item.

Any single-minded setting can be represented as a partially-ordered setting: the set of possible interests $\mathcal{G}$ is just the set of possible bundles, $2^{[m]}$. The relation is set containment: an interest G dominates an interest G', that is, $G \succ G'$, if $G \supset G'$. The distribution F is identical.

Any partially-ordered setting can be represented as a single-minded setting: we can invent items such that every interest G maps to some subset of items. For any minimal interest G (that is, G which does not dominate any other interests), map G to a new item i: $B(G) = \{i\}$. For each successive interest $G' \in N^+(G)$, map G' to $B(G') = \{j\} \cup \bigcup_{G'':G' \in N^+(G'')} B(G'')$ where j is a new additional item. Repeat this process, completing a mapping from interests to subsets of some m created items. For all subsets $B \in 2^{[m]}$ which do not have an interest that maps onto it, assign measure 0 to the event of drawing (v, B) from F. Otherwise, $f_{B(G)}(v) = f_G(v)$.

4.11 The Master Theorem

All of the analysis in the previous section started from a candidate dual solution, and showed that such duals are optimal (as in, there is a feasible primal satisfying complementary slackness). The missing step is ensuring that there exists an input distribution for which these duals are feasible. To save ourselves (and future work) the tedium of hand-crafting an actual distribution for which these duals are feasible, we prove a general Master Theorem, essentially stating that for a wide class of duals (essentially, anything dictated by ironed intervals, positive/negative regions, and flow in), there exists a distribution for which this dual is feasible.

Theorem 14 (Master Theorem)**.** *Suppose we are given a partial order over $\mathcal{G}$, for each item $G \in \mathcal{G}$ candidate endpoints of zero region (bounded away from 0) $\bar{r}_G, \underline{r}_G$, a finite set of candidate ironed intervals (bounded away from zero) $[\underline{x}_{i,G}, \overline{x}_{i,G}]$ with $\underline{r}_G \leq \underline{x}_{i,G} \leq \overline{x}_{i,G} \leq \bar{r}_G$, and for each pair of items $G' \succ G$ a finite set of candidate flow-exchanging points (bounded away from zero) $y_{i,G,G'}$* not in $(\underline{x}_{i,G}, \overline{x}_{i,G}]$ *for any candidate ironed interval. Then there exists a joint distribution over (value, interest) pairs with a feasible dual (λ, α) such that:*

- *the endpoints of the zero region for $\Phi_G^{\lambda,\alpha}$ are $\underline{r}_G$ and $\bar{r}_G$.*
- *the ironed intervals of $\Phi_G^{\lambda,\alpha}$ are exactly to the intervals $[\underline{x}_{i,G}, \overline{x}_{i,G}]$ (no others).*
- *$\alpha_{G,G'}(y) > 0 \Leftrightarrow y = y_{i,G,G'}$ for some i.*

Note that from the proof the Master Theorem, it is clear how to explicitly construct a distribution for the lower bound (although this is a tedious and unilluminating process).

In this section we provide a complete proof of Theorem 14. On our way to prove this theorem, we generalize a result of Saxena et al. [2018], in which they show that for totally ordered preferences, one can always find a discrete distribution that produces a well-enough-behaved revenue curve. They use this result to show that there exist instances for which the menu complexity is the worst possible, exponential in the number of items. Here we extend their construction and show that for any well-enough-behaved set of continuous revenue curves for the partially ordered setting, there exist distributions that induce them.

The first step is to generalize the result of Saxena et al. [2018] from discrete distributions to continuous distributions.

Lemma 9 (Revenue Theorem for Continuous Curves). *Given a continuous curve* $R : [1, H]$ *differentiable everywhere except at countably many points, such that* $R(1) = 1$ *and* $|R'(x)_+|, |R'(x)_-| \leq \frac{1}{2H} \forall x \in [1, H]$*, there exists a distribution* $\mathcal{F}$ *such that* R *is the revenue curve that arises from selling to a single bidder with a valuation drawn from* $\mathcal{F}$*.*

Proof. Consider the following distribution

$$F(x) = 1 - \frac{R(x)}{x}, x \in [1, H]$$

and $F(x) = 0$ for $x \leq 1$, $F(x) = 1$ for $x \geq H$. In order to show that this is a valid distribution, it suffices to show that it is monotonic non-decreasing. For that, we consider its derivative and show it is non-negative everywhere:

$$F'(x) = \frac{-xR'(x) + R(x)}{x^2}.$$

It suffices to show that the numerator, $R(x) - R'(x)x$, is always non-negative. Note that for $x \geq 1$, $R(x) \geq \frac{1}{2}$ (since $R(1) = 1$ and the derivative doesn't change fast enough) and $|R'(x)_+| \leq \frac{1}{2H}$. Since $x \leq H$, the claim follows.

It remains to show that indeed the revenue from this distribution matches the curve $R(x)$. Consider setting a price of x, then the revenue of selling at x is exactly $x(1 - F(x)) = R(x)$. □

Now we want to extend this to say we can find distributions for revenue curves with specific properties that will be useful.

Theorem 15 (Master Theorem for Single Item). *Given candidate endpoints of zero region* x_+, x_- *and candidate ironed interval endpoints* $[\underline{x}_i, \overline{x}_i]_{i=1}^k$ *(where* $x_- \leq \underline{x}_i \leq \overline{x}_i \leq x_+$*) there is a distribution* $\mathcal{F}$ *such that the revenue curve induced by a bidder whose valuation is drawn from* $\mathcal{F}$ *satisfies*

- $\Phi^{\lambda,\alpha}(x)f(x)$ *is negative for* $x < x_-$ *(i.e.* x_- *is the lower endpoint of the zero region),*
- $\Phi^{\lambda,\alpha}(x)f(x)$ *is positive for* $x > x_+$ *(i.e.* x_+ *is the upper endpoint of the zero region) and,*
- *the ironed intervals correspond exactly to the intervals* $[\underline{x}_i, \overline{x}_i]$ *for* $i = 1$ *to* k*.*

Proof. We will reduce the problem of finding a valid distribution to that of constructing a revenue curve that will guarantee these properties and then apply Lemma 9. Consider the following revenue curve:

$$R(x) = \begin{cases} x & 0 \le x \le 1, \\ 1 + \frac{x}{2H} & 1 \le x \le x_-, \\ 1 + \frac{x_-}{2H} & x_- \le x \le \underline{x}_1 \\ 1 + \frac{x_- + \underline{x}_1 - x}{2H} & \underline{x}_1 \le x \le \frac{\underline{x}_1 + \overline{x}_1}{2} \\ 1 + \frac{x_- + x - \overline{x}_1}{2H} & \frac{\underline{x}_1 + \overline{x}_1}{2} \le x \le \overline{x}_1 \\ \dots & \\ 1 + \frac{x_-}{2H} & \overline{x}_{i-1} \le x \le \underline{x}_i \\ 1 + \frac{x_- + \underline{x}_i - x}{2H} & \underline{x}_i \le x \le \frac{\underline{x}_i + \overline{x}_i}{2} \\ 1 + \frac{x_- + x - \overline{x}_i}{2H} & \frac{\underline{x}_i + \overline{x}_i}{2} \le x \le \overline{x}_i \\ \dots & \\ 1 + \frac{x_-}{2H} & \overline{x}_k \le x \le x_+ \\ 1 + \frac{x_-}{2H} - \frac{x - x_+}{2H(H - x_+)}(x_- + 1) & x_+ \le x \le H. \end{cases}$$

This revenue curve is such that $R(1) = 1$ and $|R'(x)| \le \frac{1}{2H}$ for $x \in [1, H]$. This allows us to claim that there is a distribution that induces this revenue curve. Moreover, from the way we constructed this revenue curve, the derivative is positive from 0 to x_-, negative from x_+ to H, goes from negative to positive for the intervals $[\underline{x}_i, \overline{x}_i]$ and is 0 elsewhere. We will show that these conditions are sufficient to make the virtual values take the signs we intend them to.

It suffices to note that the sign of the derivative of the revenue at x is the opposite of the sign of the virtual value at x (noted in Definition 5). By construction, our revenue curve has negative slope on values higher than x_+ and positive slope on points below x_-. The intervals in between will be ironed and turn into 0 slope intervals. □

Remark 16. *It is possible to relax the condition that all ironed intervals are between x_-, x_+. It is not hard to see how to adapt the proof to have ironed intervals either in $[1, x_-]$ or $[x_+, H]$. It is sufficient to add dimpled intervals, like the ones in our construction, as the revenue curve is increasing or decreasing. We don't need them for our main result, hence don't worry* *about this more general result. Likewise, the revenue curve R could be made differentiable everywhere if we used a smoother function to transition between the ironed and non-ironed intervals, as opposed to straight lines.*

Proof of Theorem 14. If the constraint over flows wasn't there, the problem would be a direct application of Theorem 15. Unfortunately, the flow constraints may affect the virtual values of neighboring items. It is not hard to predict how outgoing and incoming flow will change the virtual values for the different items. From the study of duality in this context we know that if there is ε-flow leaving from (y_i, G) to (y_i, G') (where $G' \in N^+(G)$), then the virtual values of all points of item G with $y \leq y_i$ will increase by ε and all points of item G' with $y \leq y_i$ will decrease by ε. Thus, given that we know what we want the revenue curves to look like after all flow has been sent, we can reverse engineer what they must look like in order to make that happen. In particular, since the flows shift the virtual values by a constant it will suffice to subtract a function whose value is 0 before y_i and becomes a line with small, negative slope at x_i (say, slope $\varepsilon = \frac{1}{2^H}$) from the "suggested" (by Theorem 15) revenue curve for item G (since these will increase by ε after the flow is sent) and add positive slope functions of the same value at x_i on item $G_{i,G}$ from its suggested revenue curve (since these will decrease by ε after the flow is sent). This is sufficient because of the connection between virtual values and revenue curves argued before: the derivative corresponds to changes in the virtual value. So for a constant change in virtual value, the matching change would be adding a linear term to the revenue curve of opposite sign. The order in which we do these changes is by processing items from leaves to the root (i.e. only process a node once all its children have been processed) and within an item G, address the flow-exchange values from smallest to largest. □

We abuse this opportunity to prove a similar result for the multi-unit pricing setting.

Theorem 17 (Master Theorem for MUP)**.** *Suppose we are given a MUP instance where the buyer can get up to n copies of an item. Let G_i for $1 \leq i \leq n$ be the item corresponding to i copies. For each item G_i we are given candidate endpoints of the zero region x_{-i}, x_{+i} and a set of candidate ironed interval endpoints $[\underline{x}_{j,i}, \overline{x}_{j,i}]_{j=1}^{k_i}$ with $x_{-i} \leq \underline{x}_{j,i} \leq \overline{x}_{j,i} \leq x_{+i}$. Moreover, for each tuple $(i, i+1)$ and $(i, i-1)$, we are given a set of candidate flow-exchanging points*

$y_{j,i,i+1}$ and $y_{j,i,i-1}$ not in $(\underline{x}_{j,i}, \overline{x}_{j,i}]$ for any candidate ironed interval. Then, there exists distributions $\mathcal{F}_G$ for all items G such that:

- *the endpoints of the zero region for G_i correspond to x_{-i}, x_{+i},*
- *the ironed intervals correspond exactly to the intervals $[\underline{x}_{j,i}, \overline{x}_{j,i}]_{j=1}^{k_i}$ (and no other),*
- *the dual of the problem is such that there $\alpha_{G_i,G_{i+1}}(y_{j,i,i+1}) \geq 0$ (i.e. there is flow sent from G_i at y_i to G_{i+1} into $y_{j,i,i+1}$ and no other flow from i to $i+1$).*
- *the dual of the problem is such that there $\alpha_{G_i,G_{i-1}}(y_{j,i,i-1}) \geq 0$ (i.e. there is flow sent from G_i at y_i to G_{i-1} into $\frac{i-1}{i} y_{j,i,i-1}$ and no other flow from i to $i-1$).*

Proof. This proof is similar to that of 14 with the exception that on the former, increasing the flow from (v, G) to (v, G') (with $G' \in N^+(G)$) by a little bit increases and decreases the virtual values below v by the same amount. This is no longer true since we are moving from $(y_{j,i,i-1}, G_i)$ to $(\frac{i-1}{i} y_{j,i,i-1}, G_{i-1})$. In this case, sending ε flow from $(y_{j,i,i-1}, G_i)$ to $(\frac{i-1}{i} y_{j,i,i-1}, G_{i-1})$ increases the virtual values below $(y_{j,i,i-1}, G_i)$ by ε but decreases the ones on the other end by only $\frac{i-1}{i}\varepsilon$. So, in order to reverse engineer the change in virtual value induced by this setting we need to add the same functions as in the proof of Theorem 14 to the revenue curve suggested for G_i and add a $\frac{i}{i-1}$-scaled version of it for the receiving item at the point $(\frac{i-1}{i} y_{j,i,i-1}, G_{i-1})$ on the revenue curve for G_{i-1}. The order in which these we do these changes is by processing items from leaves to root (i.e. from G_n to G_1) and within a item G_i, address the flow-exchange points from smallest to largest. □

4.12 A Candidate Dual for a Lower Bound on Menu-Complexity for the Multi-Unit Pricing Problem

Consider an MUP instance where the buyer can get one, two, or three copies of a given item. The relevant complementary slackness constraints in this setting go from

- **Rightwards.** For all v, from $(v, 1) \to (v, 2)$ and $(v, 2) \to (v, 3)$. This is because a buyer can always misreport and get more items.

- **Leftwards.** For all v, from $(v, 2) \to (v/2, 1)$ and $(v, 3) \to (2v/3, 2)$. This is because a buyer would prefer getting fewer items if they are available for much cheaper.

As shown in [DHP '17], a buyer of type (v, C)'s utility for reporting $(v/2, A)$ is given by $u_A(v/2) = \int_0^{v/2} a_A(x)dx$. The same buyer's utility for reporting (v, B) is given by $u_B(v) = 2\int_0^{v} a_B(x)dx$.

To construct a lower bound for the MUP instance, we adapt our construction from the partially ordered case. We describe our construction formally below, but note here all the relevant differences. Observe that the incentive compatibility constraints for the MUP instance described above hide a partially ordered instance inside them. Indeed, the 'item' 2 is analogous to the item C, while the items A and B are the items 1 and 3 respectively. Just like the partially ordered instance, there are incentive compatibility constraints from $(v, 2) \to (v, 3)$ for all v. The only difference is that the constraints from $(v, 2) \to (v, 1)$ have been replaced by those from $(v, 2) \to (v/2, 1)$. Also, there are 'new' constraints from $(v, 1) \to (v, 2)$ and $(v, 3) \to (2v/3, 2)$.

We claim that, despite these changes, the essence of our argument there still holds. Roughly speaking, our argument there involved constructing a top-chain (see Definition 9) oscillating between items A and B. For any value x in this chain, we had flow coming from C to *both* A and B. Reasoning about complementary slackness constraints, then, gave us our lower bound.

For the MUP case, we can still do all the above things with the caveat that the value $(v/2, 1)$ has to be treated as if it were $(v, 1)$. An analogous master theorem can still be proved as the effect of the 'diagonal' flow on the virtual values is predictable. Using the master theorem, we can construct (essentially) any dual we want. Thus, we can have a feasible dual with a top-chain of an arbitrary length M oscillating between items 1 and 3. Also, we have flow from the item 2 to both 1 and 3 at all values in this top-chain. Chasing through the complementary slackness constraints in this dual again gives us a lower bound.

To highlight this analogy, in what follows, we use C instead of 2, A instead of 1, and B instead of 3.

Formally, we construct given an integer $M > 0$, a dual containing a top chain among A and B of length M. That is, a sequence of points $(x_1, A), (x_2, B), \dots, (x_M, A)$ such that

$$\underline{x_M/2}_A < \underline{x_{M-1}}_B/2 < \cdots < \underline{x_2}_B/2 < \underline{x_1/2}_A.$$

In this dual, we have no extra space between the ironed intervals:

- $\underline{r}_A = \underline{x_M}_A/2$, $\overline{r}_A = \overline{x_1}_A/2$, and for i such that (x_i, A) and (x_{i+2}, A) are in the chain, $\underline{x_i/2}_A = \overline{x_{i+2}/2}_A$.

- $\underline{r}_B = \underline{x_M}_B$, $\overline{r}_B = \overline{x_2}_B$, and for i such that (x_i, B) and (x_{i+2}, B) are in the chain, $\underline{x_i}_B = \overline{x_{i+2}}_B$.

Recall that by definition of the $\bar{\cdot}$ and $\underline{\cdot}$ operators, $(\underline{x}_G, \overline{x}_G]$ is ironed in G. Also by our definitions, $f_G(v)\Phi_G^{\lambda,\alpha}(v) > 0$ for $v \geq \overline{r}_G$; $f_G(v)\Phi_G^{\lambda,\alpha}(v) = 0$ for $v \in [\underline{r}_G, \overline{r}_G]$; $f_G(v)\Phi_G^{\lambda,\alpha}(v) < 0$ for $v \leq \underline{r}_G$.

We will also define C to be DMR (and thus have no ironed intervals) with $\underline{r}_C = 2\underline{r}_A$ and $\overline{r}_C = 2\overline{r}_A$.

We adapt the flow from the partially ordered lower bound example: for any (x, G) in the chain, $\alpha_{C,A}(x \to x/2) > 0$ and $\alpha_{C,B}(x \to x) > 0$.

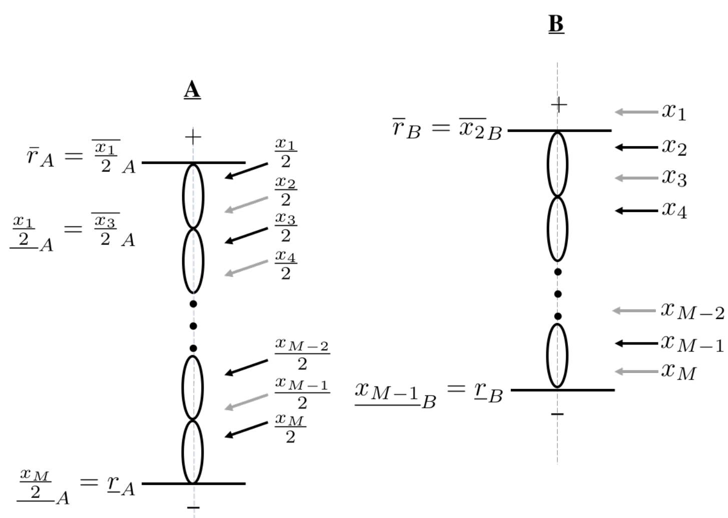


Figure 4.19: The analogue of the partially ordered candidate dual, adjusted for the Multi-Unit Pricing problem.

Theorem 18. *To satisfy complementary slackness with the candidate dual, the allocation requires M distinct allocation probabilities; the menu complexity is at least M.*

Proof. The proof is almost identical to that of Theorem 3. Using the constraint that the allocation can't increase in the middle of an ironed interval and that $u_A(x/2) = u_B(x)$

for all (x, G) in the chain, we show that the allocations must be non-zero throughout the chain.

Then, we show that for consecutive points in a chain $(x_i, A), (x_{i+1}, B)$ that $(1/2)a_A(x_i/2) > 2a_B(x_{i+1})$, and similarly, for $(x_i, B), (x_{i+1}, A)$, that $2a_B(x_i) > (1/2)a_A(x_{i+1}/2)$

This is enough to show that all of the menu options must be distinct, requiring meu complexity $\geq M$. □

Part II

Proportional Complementarities

5 Proportional Complementarities

5.1 Introduction

Consider a setting where multiple items are being sold, and a buyer's valuations for the items have *complementarities*. That is, the buyer derives some value from owning a combination of items that is not present when owning any of the items individually, as in the following examples.

Microsoft Office Example: A person who values producing documents will value software such as Microsoft Word that helps him in this task. If the person wants to include some charts in his document, then this is made easier and faster by having another piece of software that specializes in making charts such as Microsoft Excel. Thus owning Excel in addition to Word boosts the value of Word for him, since he can then produce more documents in the same amount of time.

Cloud Services Example: A cloud service provider offers multiple heterogeneous items that are both substitutes and complements. You can purchase a general purpose virtual machine (VM) or a special purpose VM such as a "data science VM"; these are substitutes. You can also purchase an upgrade such as a fast solid state disk-drive (SSD) which would be complementary to either of those VMs.

The goal of this work is to understand how a revenue-maximizing seller should price items such as Microsoft Office products or cloud services when facing a buyer with such complementarities. To this end, we introduce a new model of complementarities, design a pricing scheme for this model, and show worst-case approximation guarantees.

5.1.1 Related Work: Simple and Approximately Optimal Pricings

In recent years, there has been a surge of research activity on *optimal combinatorial pricing*. This is the problem of determining and pricing bundles of heterogeneous

This chapter is based on joint work with Yang Cai, Nikhil Devanur, and Preston McAfee in a paper titled "Simple and Approximately Optimal Pricing for Proportional Complementarities" which appeared at EC 2019 [CDGM '19].

items in order to maximize revenue from selling to a buyer who has a combinatorial valuation function. The theme of the research has been *simple vs. optimal*, where simple pricing schemes are shown to approximate the optimal (possibly randomized) pricing scheme to within a universal constant multiplicative factor, independent of the number of items.

One of the first such results was by Chawla et al. [2010], who consider the setting with n *unit-demand* buyers, who each want at most one item, with values drawn independently across n buyers and m items. They prove that selling the items separately, posting the Myerson reserve price on each item, gives a 30-approximation to the optimal revenue. They do this by defining a "copies" setting, where each buyer i is represented by m single-dimensional agents, one for each item j that buyer i might be interested in; further, only one of the m agents representing i can be served, just as agent i can only receive one item in the unit-demand setting. However, because the m agents for buyer i compete with each other in the copies setting, whereas they cooperate in the unit-demand setting, the optimal revenue from the copies setting is higher than the optimal revenue from the unit-demand setting (but this is just intuition). Then, Chawla et al. use a prophet-inequality-like argument to set prices in the unit-demand setting based on the optimal prices in the copies setting, earning a constant-fraction of the optimal revenue from the copies setting. By setting the prices to sell with half the ex-ante probability that they did originally, it ensures that each item's price is high enough, while each item is bought frequently enough.

The second such result is for a single *additive* buyer with values drawn independently across m items. Babaioff, Immorlica, Lucier, and Weinberg [2014] show that the better of (1) selling separately, or posting the Myerson reserve price on each item, and (2) selling the grand bundle, or posting the reserve pricing for the distribution of the sum of all item values together, gives a 6-approximation to the optimal revenue in expectation. This is surprising, since neither selling separately nor grand bundling on their own can guarantee a constant-factor approximation [Hart and Nisan, 2017]. Babaioff et al. perform a "core-tail analysis" on the probability distributions, essentially partitioning the event space into the "tail," when the buyer has an extremely high value for any item, and the "core," the rest of the time. They show that the optimal revenue is bounded by the revenue of the tail and the welfare of the core. Then, they show that selling separately earns a constant fraction of the revenue of the tail. Finally, either selling separately also earns a constant fraction of the welfare of the core, or the core must concentrate, and thus, selling the grand bundle earns a constant fraction of the welfare of the core.

Next, Rubinstein and Weinberg [2015] extend an analogue of this result for a single subadditive buyer, yielding a 338-approximation to the optimal revenue. They relax the buyer's valuation function to one that is additive between the set of items in the tail and those in the core, creating a valuation function that is only larger. Then, they show that *approximate revenue monotonicity* holds: if a distribution F_H first order stochastically dominates F_L, then the optimal revenue from a buyer with valuations drawn from F_H is within a constant factor of that from a buyer with valuations drawn from F_L. (It is known that in such settings, strict revenue monotonicity does not hold [Hart and Reny, 2012].) These pieces allowed the authors to finish off an analysis similar to that of Babaioff et al..

Yao [2015] produced the first analogue of Babaioff et al. for multiple additive bidders: the better of selling separately and a "two-part tariff" (an entry fee plus item prices) yields a 57-approximation to the optimal revenue. He uses a technique from the Lookahead Auction [Ronen, 2001] to ensure that only the highest bidder for each item competes for it. The idea is to force each buyer i to pay for item j at least the maximum value of any other bidder for item j, β_{ij}, and thus only the highest bidder will be able to afford the item, ensuring that at most one buyer buys each item. Then, as this is the baseline price, all of the distributions F_{ij} are shifted down by β_{ij} (with all negative values pushed up to zero) and the single-agent problem from Babaioff et al. is solved. As the β's are added back in as item prices, the "bundle" pricing becomes an entry fee for the right to buy any items, and then buyer i can take any remaining item at a price of β_{ij}, or the "selling separately" item prices are added onto the β's to form higher item prices.

The next extension in this line of work was was by Chawla and Miller [2016] for multiple constrained-additive buyers: the better of selling separately and a sequential two-part tariff yields a 133-approximation. Each buyer can be additive subject to a publicly known matroid feasibility constraint, and these feasibility constraints need not be identical among buyers. As a result, this is the first work to handle a population comprised of both additive and unit-demand buyers (and other constrained-additive buyers as well). Their approach is to solve for the optimal ex-ante sale probabilities for each buyer and item, and then to solve the single-buyer problem subject to these ex-ante constraints. Then, they stitch together the single-buyer solutions into a sequential mechanism.

Following this, Cai et al. [2016] put forth a Lagrangian duality framework that managed to unify the approximation results of Chawla et al. [2010], Babaioff et al. [2014], and Yao [2015], despite the differences in valuations and numbers of bidders. The approach is to formulate the (primal) optimization problem of maximizing expected revenue subject to incentive-compatibility (IC) and feasibility, and then to take the partial Lagrangian by multiplying the IC constraints by Lagrangian multipliers and moving them into the objective. Then, when the dual of this problem has its variables set optimally, its objective is equal to that of the primal under strong duality, and for any feasible variables the dual provides an upper bound. Cai et al. observe that the dual only provides a finite upper bound when the payment variables satisfy an equation that they interpret as a "flow constraint," and thus, they interpret the rest of the dual variables as forming a flow. They construct "canonical" dual variables that are analogous to the single-dimensional optimal dual variables, and call the resulting upper bound on the optimal revenue the *benchmark*. The benchmark divides into several terms, which the authors then bound using the optimal revenue in the copies setting or the core-tail analysis. They eventually bound these terms with selling separately or grand bundling on shifted distributions, as in prior works.

Cai and Zhao [2017] use this same Lagrangian duality framework, extending the techniques to provide a 268-approximation for multiple XOS buyers and an $O(\log m)$-approximation for multiple subadditive buyers. (Their approach also improves many of the previous constants, such as the 338 from [Rubinstein and Weinberg, 2015] to 40.) As with [Rubinstein and Weinberg, 2015], they use a valuation relaxation tactic in order to make the duality technology work and produce a viable upper bound on the optimal revenue; this is the first benchmark produced for multiple subadditive buyers with or without duality. The bounding of the benchmark is more involved in this case, involving a double core-tail analysis and shifted distributions within the core. While this first appears similar to the approaches in Yao [2015] and Chawla and Miller [2016], this approach instead uses a prophet-inequality-like analysis in combination with a concentration inequality for subadditive functions.

5.1.2 Related Work: Complementary Valuations

All of the above simple-and-approximately-optimal results apply only in settings where, at minimum, subadditivity holds. This assumes that a buyer's valuations are complement-free. However, there are many settings where a buyer derives some extra value from owning a combination of items that is not present from owning any item individually; that is, the items have *complementarities*. Further, in practice, bundling is most attractive when the items are complementary to each other.

Complementarities in valuations were, for a long time, notoriously difficult not only for revenue, but for welfare maximization in polynomial time as well[1]. While complement-free valuations admit constant-factor polynomial time approximations [Feige, 2009], or even allow VCG to be run in polynomial time, this is far from true for single-minded valuations, for which the best approximation in polynomial time is $O(\sqrt{m})$ for m items [Lehmann, O'Callaghan, and Shoham, 2002].

The first major progress was made by Abraham et al. [2012] who suggested a restricted model of complements that came to be known as the positive-hypergraph-k (PH-k) model. A hypergraph captures the valuations, where each vertex in the hypergraph corresponds to an item, and the vertex i has weight equal to the buyer's value for item i alone, w_i. Then, for any subset of items T with size at most k, there may exist a weighted edge w_T in the hypergraph corresponding to the value of some activity that can be done with that collection of items. A buyer's valuation for any set of items S is then the sum over his value for the activities associated with all subsets of items $T \subseteq S$, that is,

$$v(S) = \sum_{i \in S} w_i + \sum_{T \subseteq S} w_T.$$

For valuations from PH-k, Abraham et al. show that algorithmically, a k-approximation to welfare can be given in polynomial time[2]. However, their truthful mechanism only obtains an approximation factor of $O(\log^k m)$.

Following this, Feige et al. [2015] defined a hierarchy of valuations that generalize the PH-k model, as well as many other models of restricted complements. They define the class of *maximum-over-positive-hypergraph-k* valuations (MPH-k). That is,

[1] Welfare maximization disregarding running time is solved by the VCG mechanism.

[2] Abraham et al. use demand queries in their algorithm, but show that for the PH-k model, demand queries can be implemented in polynomial time.

there exist valuation functions $\{v_\ell\}_{\ell \in \mathcal{L}}$ such that $v_\ell \in$PH-$k\ \forall \ell$ and $v(S) = \max_{\ell \in \mathcal{L}} v_\ell(S)$. First, they show the expressiveness of this hierarchy of valuations to capture existing classes. In particular, MPH-1 is equivalent to the class of XOS valuations (maximum over additive valuations), and contains submodular valuations. Every monotone graphical function (as defined in [Conitzer, Sandholm, and Santi, 2005]) is contained in MPH-2. The authors show that the k^{th} level of the XOS hierarchy is contained in MPH-k, as is every function with supermodular degree k (as defined in [Feige and Izsak, 2013]), and clearly PH-$k \subset$ MPH-k. Interestingly, however, the converse directions do not hold, demonstrating that the MPH hierarchy is strictly more expressive. Further, Feige et al. prove that, assuming access to demand oracles, they can obtain an algorithmic $(k+1)$-approximation to welfare in polynomial time for valuations in MPH-k, essentially matching the result of Abraham et al. [2012]. Finally, they show that simultaneous first-price auctions guarantee a price of anarchy of at most $2k$ for bidders with MPH-k valuations, a much stronger result than Abraham et al. when $k \ll m$, although simultaneous first-price auctions are of course not truthful. Their results extend for Bayes-Nash and correlated equilibria, assuming the distributions are independent across buyers.

The first work focusing on revenue maximization for complements is that of Eden et al. [2017b], who aim to find a simple and approximately-optimal mechanism for a single buyer with complementarities from the MPH-k model. They assume the buyer's hyperedge types are drawn independently from known distributions. Since the valuations are additive over hyperedges, the authors apply the [CDW '16] framework for additive valuations to the hyperedges. This implies that the better of selling the hyperedges separately and grand bundling them is a good approximation to the optimal revenue. Since grand bundling the hyperedges is the same as selling the grand bundle of items, it remains to understand how to approximate selling the hyperedges separately by selling items. First, they approximate the separate hyperedges with separate *disjoint* hyperedges, losing a factor of d, and then they sell a specific item for each hyperedge (which are now disjoint) and give the remaining items away for free. The result is that the better of selling separately and the grand bundle in the MPH model gives an $O(d)$-approximation, where d is the largest degree of any vertex, or the largest number of edges that any one item appears in. Further, they prove that this factor d is necessary. In addition, they show that the approximation would be exponential in the positive rank k, despite the fact that the welfare approximations from Abraham et al. and Feige et al. are in k.

5.1.3 Proportional Complementarities Model

The PH valuation model for the Microsoft Office example would have three values, one each for Word, Excel, and the pair (Word, Excel), with each of them drawn independently from a different distribution. While having the values for Word and Excel be independent may be reasonable, that the value for the pair (Word, Excel) be independent of the other two seems unrealistic. Similarly, for the cloud services example, the PH model would have that the value for the pair (VM, SSD) be independent of the value for the VM alone, which is once again unrealistic.

We introduce what we call a *proportional complementarities* model of valuations; a special case of this model is *proportional pairwise complementarties* (PPC). We illustrate this model through the examples we considered before.

Microsoft Office Example in the PPC model: We still have a value for each of Word and Excel, say v_1 and v_2 respectively, that are independent of each other. Our model differs in how the buyer values the combination of the two by assuming that the additional value derived from having both items is due to a better utilization of either item, and hence is proportional to (rather than independent of) the buyer's base valuation for Word and for Excel. This is captured in our model by having a multiplier for the pair (Word, Excel), denoted by $\eta_{1,2}$; say Excel always adds 23% to the value of Word, then we would have $\eta_{1,2} = 0.23$. One could get an estimate of this quantity by observing the frequency of activities between the two, such as dragging Excel charts into Word. While not fully general, these proportionalities make intuitive sense, because if a buyer values an item highly, he is likely to care more about its complements too, as they enhance that item. The value for purchasing both items in our model would then be

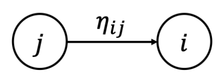
$$v_1(1 + \eta_{1,2}) + v_2.$$

The other assumption we make is that while the seller doesn't know the exact values, he knows these proportions of complementarities. This is perhaps the least accurate assumption in applications, because such values could reasonably vary across individuals. However, in circumstances where the way products are used together is approximately fixed, such as dragging Excel charts into Word, it is not unreasonable to assume that these values are known. This is especially true when it comes to "digital goods," where data about interactions between items can be gathered, and the parameters can be estimated from this data, e.g. via estimating cross-price elasticities.

Allowing the proportions (i.e., the ηs) to vary across individuals is an interesting direction for future research. We present one possible approach via a common generalization of our model and the PH model in Section 5.5. This generalization further illustrate the similarities and the differences between the two models.

Proportional pairwise complementarities: We first define the PPC model. A single seller offers m heterogeneous items for sale to a single buyer. (Equivalently, there is a population of buyers, but no supply constraints on the seller, as is the case with digital goods like Microsoft Office products.) We model the structure of the complementarities among the items via the following parameters, which are assumed to be known to the seller:[3]

$$\eta_{ij} \in \mathbb{R}_+ \qquad \forall\, i, j \in [m], i \neq j.$$



The parameter η_{ij} captures how much having item j boosts the valuation that the buyer derives from item i. The valuation of a buyer is determined by his type t, which is a vector in $\mathbb{R}_+^m$, and is the private information of the buyer. The i^{th} coordinate of t is t_i, which represents his base valuation for item i in the absence of any other items. If the buyer also gets item j, then his valuation for item i is boosted by an additional $\eta_{ij} t_i$. From this, we get that for any bundle $S \subseteq [m]$, the buyer's valuation for S is

$$v(t, S) := \sum_{i \in S} \eta_i(S) t_i, \qquad \text{where} \qquad \eta_i(S) = 1 + \sum_{j \in S \setminus \{i\}} \eta_{ij}.$$

Note that η_{ij} need not be equal to η_{ji}, and asymmetric boosts are only more general. We make the Bayesian assumption that t is drawn from a product distribution $\Pi_{i \in m} F_i$.

[3]We use the notation $[m]$ to indicate the set of first m natural numbers, $\{1, 2, \ldots, m\}$.

The distributions F_i for all $i \in [m]$ (as well as the parameters η_{ij}) are known to the seller.

This more general asymmetric case corresponds to directed graphs (and hypergraphs). Thus we define the *directed-positive-rank* k of the graph to be the maximum size of (number of items in) the *source* of a (hyper)edge. Thus, for the pairwise case, $k = 1$.

The general case: The general class of valuations we consider is defined formally in Section 5.2; we give an informal description here. First of all, we allow hyperedges, instead of edges, i.e., each pair of item i and a disjoint set of items T forms a directed hyperedge (T, i) and has a certain boost associated with it, denoted by η_{iT}: this is the boost of having all items in T on item i. The valuation of a set S now includes all possible boosts due to hyperedges (T, i) for $T \sqcup \{i\} \subseteq S$. We call this class of valuations *proportional positive hypergraphic (PPH) valuations*. The other generalization is to allow the boost to be the maximum of the boost from multiple hypergraphs. We call this class of valuations *maximum of proportional positive hypergraphic (MPPH) valuations*. We denote by k the directed-positive-rank and by d the maximum-degree of the hypergraph. We tie this back to the cloud services example to show how such a generalization is useful.

Cloud Services Example: Suppose that we had access to two types of VMs, VM1 and VM2, that are meant for different types of workloads. We can also purchase additional disk drives (DDs) that allow us to run larger workloads. DDs come in two technologies, fast and slow, say DD1 and DD2. Having either of the DDs can boost the value for a VM, and having both of them boosts it even more but less than the sum of the individual boosts. This could be modeled as follows. There are 4 items, 1 and 2 are the VMs, and 3 and 4 are the DDs. For each of $i \in \{1, 2\}$ and $j \in \{3, 4\}$, we have the boosts η_{ij} as well as $\eta_{i\{3,4\}}$. Let x_3 and x_4 be binary variables indicating whether items 3 and 4 were respectively purchased or not. The value derived from item i for $i \in \{1, 2\}$ depending on these choices is

$$t_i \cdot (1 + \max\{\eta_{i3}x_3, \eta_{i4}x_4, \eta_{i\{3,4\}}x_3x_4\}).$$

Thus VM1 can get a boost of η_{13} from having DD1, or η_{14} from DD2, but if you have both DD1 and DD2, the boost is $\eta_{1\{3,4\}}$ rather than $\eta_{13} + \eta_{14}$.

5.1.4 Pricing scheme

Almost all of the papers in this line of research consider the better of selling each item separately and selling only the grand bundle. Pricing the grand bundle is (conceptually) easy: set the monopoly price for the distribution of the buyer's value for the grand bundle, which can be computed from the given input. For simple valuations such as additive valuations, setting item prices to sell separately is also easy: set the monopoly reserve for each of them separately. In our model, this completely ignores the boost in the valuation on an item from having other items. Not surprisingly, this can be provably far from optimum when you have complementarities; we therefore need a non-trivial way to price the items in this case. We first illustrate our algorithm for finding these prices via a numerical example.

Numerical Example: Suppose, as shown on the left in Figure 5.1, that there are 4 items, numbered 1 through 4, and that we have non-zero ηs on the pairs $(2, 1)$, $(3, 2)$, $(4, 3)$ and $(1, 4)$. Let all ηs be 1. Suppose t_1 and t_3 are distributed identically as follows: the value is 2 w.p. $\frac{1}{2}$ and 0 otherwise; let t_2 and t_4 be distributed identically as follows: the value is 4 w.p. $\frac{1}{2}$ and 0 otherwise. Each t_i is independent of the others.

We denote the monopoly price and the monopoly revenue for item i alone by r_i and R_i respectively. For this example, we have the monopoly prices as $r_1 = r_3 = 2$ and $r_2 = r_4 = 4$; the revenues are $R_1 = R_3 = 1$ and $R_2 = R_4 = 2$. Setting the monopoly prices for each item separately guarantees a revenue of $\sum_i R_i = 6$. The actual revenue would be higher, but in general it is difficult to get a better handle on it than this bound.

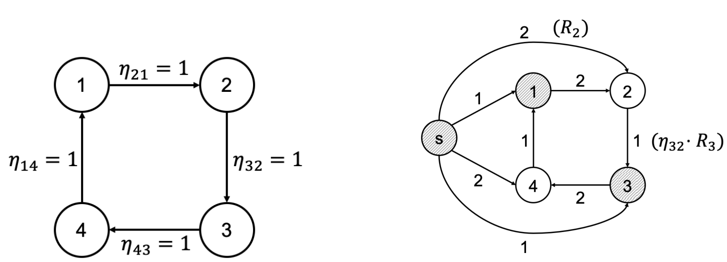


Figure 5.1: Left: The 4-item example described above. Right: The directed graph where the weight of a directed cut corresponds to a lower bound on the revenue of the corresponding SEPARATE/FREE mechanism.

Step 1: Construct a weighted directed graph. We construct a weighted directed graph with 5 vertices, one for each item, and a source node s. The weight on the edge (s, i) is R_i. The weight on the edge (i, j) is $\eta_{ji} R_j$. This graph is shown on the right in Figure 5.1.

Step 2: Find a max directed cut. We then find a cut in the graph that maximizes the number of directed edges going from the "source" side to the "sink" side. From the figure, it is easy to see that such a cut is given by the vertices $s, 1$ and 3 on the source side, the rest on the sink side, and has weight 8.

Step 3: Set Prices. We set the prices for items on each side of the cut differently.

1. The items on the source side have a price of 0. This set of items, denoted by $\mathcal{F}$, are "free". In this case, items 1 and 3 are free.
2. For the items on the sink side, we multiply the monopoly price r_i by $\eta_i(\mathcal{F})$ (1 plus the boost i gets from all the items on the source side). Then items 2 and 4 thus have a price of 8 each.

The weight of the cut, 8, is a lower bound on the revenue of this pricing scheme. Each of items 2 and 4 is bought at the price of 8 with probability $\frac{1}{2}$, giving a lower bound on revenue of 8. In comparison, the best price for grand bundling is 12, which is bought with probability $\frac{5}{8}$, giving a revenue of $\frac{15}{2}$, which is slightly lower. Both of these are still higher than the revenue lower bound of 6 from setting separate prices of r_i each.

In general, we introduce a class of mechanisms which we call SEPARATE/FREE. Like selling separately, every item is sold separately at some price, and the buyer may take any set of items so long as he pays the sum of their individual prices. However, we partition the items into "free items" $\mathcal{F}$, where for each item $i \in \mathcal{F}$, the individual price of each such item is $\$0$, and "priced items" $\bar{\mathcal{F}} = [m] \setminus \mathcal{F}$. Once the free set $\mathcal{F}$ is determined, we use the knowledge that the buyer will take the free items to inflate the monopoly prices of the priced items by the boost on the item from also getting the free set (and only the free set).

Such mechanisms do capture a certain economic intuition that is seen in practice: giving some items away for free in order to charge more for complementary items, e.g., Google sells the Android OS for free since it is complementary to advertising revenue. One can also think of it as a certain form of bundling: there is no reason to give away the free items unless the buyer purchases some priced item. This is equivalent to bundling all the free items with any non-empty subset of paid items. Going back to our cloud services example, such a pricing scheme could determine that one of the two DDs should be free. We would then bundle that DD into the VMs; such bundles are commonly observed in practice.

One difficulty in the above scheme is that in general, finding a max directed cut in a graph is an *NP-Hard* problem. When restricted to polynomial time algorithms, the best worst-case approximation guarantee we can show is by placing each item independently into the free set with some probability α, which is determined by $\min\{d, k\}$. This is a little unsatisfactory since it doesn't use the specific market parameters η at all. (However, they are used in setting the prices once $\mathcal{F}$ is determined.) An alternative is to use an approximation algorithm for the max directed cut problem, such as the Goemans-Williamson algorithm. The advantage of this method is that it produces a free set that makes use of the structure of the ηs; unfortunately, this does not improve the worst case approximation ratio. In fact, we show that no algorithm can improve the approximation ratio when used in conjunction with our current proof technique, but we conjecture that such an algorithm would be better in practice.

5.1.5 Worst case approximation guarantee

Once again, we begin by illustrating our analysis using the numerical example earlier. For the sake of analysis, we consider an instance of the pricing problem on the same set of items, with additive valuations. The value distribution for item i in this instance, denoted by $\hat{F}_i$, is just the original distribution F_i multiplied by $\eta_i([m])$, the boost i can obtain from all of the items. In our example, $\hat{F}_i$ for $i = 1$ and 3 is 4 w.p. $\frac{1}{2}$ and 0 otherwise, and for $i = 2$ and 4 is 8 w.p. $\frac{1}{2}$ and 0 otherwise.

We relate the revenue from selling separately and selling the grand bundle on the given instance to the corresponding mechanisms for the additive instance. It is easy to see that the bundle revenue remains the same in both instances, as the complements buyer receives the boosts $\eta_i([m])$ on every item:

$$\textsc{brev-ppc}(F) = \textsc{brev-additive}(\hat{F}).$$

As we computed earlier, a lower bound on selling separately with our pricing scheme for the given instance is 8. Selling separately for the additive instance gives a revenue of 12, which is 3/2 times 8. So for this example, we have that

$$\text{SREV-PPC}(F) \geq \frac{2}{3}\text{SREV-ADDITIVE}(\hat{F}).$$

More generally, SREV-ADDITIVE$(\hat{F})$ is equal to the total the weight of all the edges in the digraph that we construct. If you place each item on either side of the cut with equal probability, then each edge is cut with probability $\frac{1}{4}$, which results in a factor of 4 between the two SREVs. This is indeed tight: consider a complete unweighted digraph; any cut can only cut a $\frac{1}{4}^{\text{th}}$ fraction of edges.

We can now use a slight generalization of the result of Babaioff et al. [2014] to bound the optimum revenue for the additive instance, denoted by OPT-ADDITIVE$(\hat{F})$, in terms of SREV and BREV.

$$\text{OPT-ADDITIVE}(\hat{F}) \leq 2\,\text{SREV-ADDITIVE}(\hat{F}) + 4\,\text{BREV-ADDITIVE}(\hat{F}).$$

Finally, we show that the optimum revenue for the additive instance is only higher.

$$\text{OPT-PPC}(F) \leq \text{OPT-ADDITIVE}(\hat{F}),$$

which gives an approximation ratio of 7 for this example, and 12 in general, working through the inequalities above. Note that even for additive valuations, 5.2 is the best known approximation ratio.

This last step may seem obvious, but it turns out to be quite tricky. One might expect a direct argument, that given a mechanism M for the original instance, we construct a mechanism M' for the additive instance, with a larger revenue. Such approaches are inherently difficult, as evidenced by "Revenue non-monotonicity" in Hart and Reny [2012]. We instead argue the upper bound by covering the dual of the smaller setting with the dual of the larger setting, a novel use of the [CDW '16] Lagrangian duality framework.

We show the following approximation guarantee more generally.

Theorem 19 (Informal)**.** *The better of* BREV *and the revenue from a mechanism of type* SEPARATE/FREE *is an* $O(\min\{d, k\})$*-factor approximation to the optimal revenue for valuations* *in the class* MPPH*. When $k = 1$, i.e., the boosts are the maximum over directed graphs, the approximation factor is at most 12.*

Recall that d is the maximum-degree of the hypergraph, and k is the directed-positive-rank of the hypergraph.

We also show that our analysis of Theorem 19 is tight up to a constant factor via the following lower bound. A crucial step in our analysis is to upper bound the optimal revenue for MPPH valuations by the optimal revenue for an instance of additive valuations. Further, the actual revenue of a mechanism from a buyer with proportional complements is extremely difficult to analyze. Instead, we analyze a lower bound on the revenue we deem the "proxy revenue," and we show that with respect to our upper bound, no mechanism of the following type can give an $o(k)$-approximation to the proxy revenue. The mechanisms we consider first partition the set of items into bundles, designating one bundle as the free set. Each of the other bundles is priced separately. The buyer always gets the free set for free. Specifically, the price for a bundle is its monopoly reserve price inflated by the boosts of only the other items in its own bundle and by the free set, and not by anything else. The proxy revenue undercounts the revenue in the same way, by assuming that the buyer's boosted values match the way prices are set in these mechanisms: only within bundles and from the free set. We elaborate on motivation for using this proxy in Section 5.3.

5.2 Preliminaries

We now give the formal description of the MPPH valuation model. There is a single seller offering m heterogeneous items for sale to a single buyer. The following parameters determine the structure of complementarities among items via boosts to base valuations. There is a hypergraph with the set of items $[m]$ as vertices whose edges (T, i) correspond to a combination of items T and a disjoint item i to which the combination gives a boost. Moreover, there could be several possible boosts out of which only the highest is activated. For each item $i \in [m]$, for each hyperedge (T, i), and for each $\ell \in [K]$ for some integer K, we have the parameter $\eta^{\ell}_{iT} \in \mathbb{R}_+$.

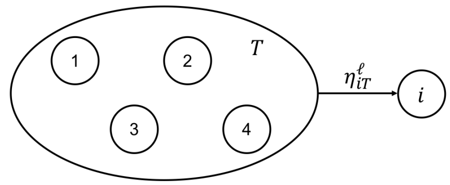


Figure 5.2: A directed graph representation of the η parameters.

The buyer's valuation for any bundle $S \subseteq [m]$ is

$$v(t,S) = \sum_{i \in S} \eta_i(S) t_i, \qquad \text{where} \qquad \eta_i(S) = 1 + \max_{\ell \in [K]} \sum_{T \subseteq S \setminus \{i\}} \eta_{iT}^{\ell}.$$

We refer to the case where the boosts are simply the sum (i.e. $K = 1$) as additive boosts, and the general case ($K > 1$) as XOS boosts. Note that $\eta_i(S)$ always includes the base valuation for item i (the $+1$) so it is not entirely comprised of boosts, but we overload and call this term the boost anyway. Observe that the boosts are always monotone in the set, that is, if $\ell(S) \in \operatorname{argmax}_{\ell \in [K]} \sum_{T \subseteq S \setminus \{i\}} \eta_{iT}^{\ell(S)}$, then it always the case that for all $S \subseteq S'$,

$$\eta_i(S) = 1 + \sum_{T \subseteq S \setminus \{i\}} \eta_{iT}^{\ell(S)} \quad \leq \quad 1 + \sum_{T \subseteq S' \setminus \{i\}} \eta_{iT}^{\ell(S)} \quad \leq \quad 1 + \max_{\ell \in [K]} \sum_{T \subseteq S' \setminus \{i\}} \eta_{iT}^{\ell} = \eta_i(S') \tag{5.1}$$

We assume that t is drawn from a product distribution $F = \Pi_{i=1}^{m} F_i$. The distributions F_i for all $i \in [m]$ and the ηs are all known to the seller. However, the type realization t is private information of the buyer.

Our approximation ratios depend on the parameters k and d of the underlying hypergraph. The parameter k, the directed-positive-rank, is an upper bound on the size of the set in any hyperedge, i.e., $|T| \leq k$ for each hyperedge (T, i). The parameter d, the maximum-out-degree, is an upper bound on the number of hyperedges that contain a particular vertex, i.e., for each $i \in [m], |\{$ hyperedge $(T, j) : i \in T\}| \leq d$. We suppress the dependence on the hypergraph in our notation, since it should always be clear from the context. For the special case of pairwise complementarities (PPC) we follow the notation in Section 5.1.3.

5.2.1 Optimal Mechanisms in Various Settings

From the revelation principle, we can restrict our attention to direct revelation mechanisms, where the buyer reports his type. A mechanism is therefore defined by the allocation and the payment functions. We allow randomized allocation rules, with the assumption that the buyer is risk neutral. Let $x_S(t)$ denote the probability that the bundle $S \subseteq [m]$ is allocated to the buyer of type t; let $p(t)$ be his payment. The incentive-compatibility (IC) constraints require that for each buyer type, the buyer maximizes utility by reporting his true type.[4] Among all IC mechanisms, the optimal mechanism maximizes the expected revenue

$$\mathbb{E}_t[p(t)].$$

Notation: We use the following convention to denote the revenue from a particular mechanism for a given class of valuations, for a particular distribution over types:

$$\text{[Mechanism name]-[Valuation Class]}([\text{Distribution}]).$$

For example, the optimal mechanism for PPC valuations with types drawn from F is denoted by OPT-PPC(F). We drop the distribution when it is clear from the context. We also drop the valuation class when it is additive (ADDITIVE) and it is clear from the context: e.g., the revenue from selling the grand bundle for additive valuations on types drawn from the distribution F is just BREV.

5.3 A Constant-Factor Approximation via a Random Free Set

We begin with the case of pairwise complementarities and show a 12-approximation for this setting.

Recall that the two standard mechanisms considered in previous work are selling the grand bundle and selling each item separately. Selling the grand bundle only gets better with complements, since we are certain that the buyer will receive all possible boosts, and we can price accordingly. It is selling the items separately that is

[4] We do not formally define IC constraints since we can bypass it due to Lemma 12, but our mechanisms will be clearly IC.

problematic. A conservative way to set the prices while selling separately is to ignore the complementarities, and sell them as if they are just additive; this could clearly be quite suboptimal. We can price an item more aggressively in order to capture some of the boost from complementarities, but this will decrease its probability of sale, which can further decrease the probabilities of sale for other items that receive a boost from this item. The pricing must get the right tradeoff between capturing more of the boost from complementarity while making sure that sufficient quantity of items are sold in the first place in order for the boosts to accrue. Overall, it is difficult to characterize the behavior of the buyer, which makes optimizing the prices extremely challenging.

Our approach is to shift the focus away from optimizing prices. We do this by giving some items away for free, and then just selling the remaining items individually as if they are additive, but accounting the boost from the items that are given for free. The free items make sure that sufficient boosts accrue; the priced items extract the value thus generated. The problem now becomes one of choosing the set of free items, but in fact we show that a random choice suffices. The analysis compares the revenue to a seemingly crude upper bound, where every item receives the fullest boost that an item could possibly receive—the boost on the item if the buyer were to receive all of the items, that is, the grand bundle.

We now formally describe our mechanism SEPARATE/FREE. For each item $i \in [m]$, let r_i^* be the monopoly reserve for the distribution F_i, i.e.,

$$r_i^* = \arg\max_{p \in \mathbb{R}_+} p \cdot (1 - F_i(p)) ,$$

and let R_i be the revenue of the monopoly reserve for the distribution F_i,

$$R_i := r_i^* \cdot (1 - F_i(r_i^*)) .$$

Mechanism SEPARATE/FREE$(\mathcal{F})$ **:** Partition the items into "free items" $\mathcal{F}$ and "priced items" $\bar{\mathcal{F}} = [m] \setminus \mathcal{F}$. The price of a priced item $i \in \bar{\mathcal{F}}$ is

$$p_i = \eta_i(\mathcal{F}) \cdot r_i^*.$$

The buyer gets all of the items in $\mathcal{F}$ for free, that is, they are priced each at 0. We denote by SEPARATE/FREE$(\mathcal{F})$ the expected revenue from the mechanism with (potentially random) free set $\mathcal{F}$, and we overload notation slightly to use SEPARATE/FREE $=$ $\max_{\mathcal{F}\subseteq[m]}$ SEPARATE/FREE$(\mathcal{F})$.

Theorem 20. *The better of selling the grand bundle and Mechanism* SEPARATE/FREE *is a 12-approximation for* PPC *valuations:*

$$\text{OPT-PPC} \leq 12 \max\{\text{BREV-PPC}, \text{SEPARATE/FREE-PPC}\}.$$

5.3.1 Proof of Theorem 20

The proof of this theorem is largely along the lines of the analysis described in Section 5.1.5. We first relate OPT-PPC to the optimal revenue for an instance of additive valuations; where the buyer's valuation for each item is inflated as if he receives the boosts from owning every possible item in addition to this one, even if he receives no additional items. Then, the buyer's new (much larger) valuations are additive. We refer to this setting as the fully-boosted additive setting, where we call the values t multiplied by the full boosts as drawn from the distribution $\hat{F}$, even though t is drawn identically as from F. We show that the revenue from this setting is only larger than from the proportional complements setting.

Lemma 10.

$$\text{OPT-PPC}(F) \leq \text{OPT-ADDITIVE}(\hat{F}).$$

This is a very loose upper bound and intuitively it should be true: for every type t, the buyer's value for every set in the fully-boosted additive setting is only larger than in the proportional complements setting. However, due to revenue non-monotonicities, the proof requires more care, and is deferred to Subsection 5.3.2.

In Appendix B.1, we prove improve the analysis of the 6-approximation by Babaioff et al. [2014] to allow a parameterization in the bound[5]. Then our Theorem 33 with $a = 1$ gives that

$$\text{OPT-ADDITIVE}(\hat{F}) \leq 2\,\text{SREV}(\hat{F}) + 4\,\text{BREV}(\hat{F}).$$

It is easy to see that the revenue from grand bundling in the complements setting on the original distribution is the same as the grand bundling in the fully-boosted

[5]This analysis also improves the 6-approximation to 5.382. The state of the art coefficient is 5.2 Ma and Simchi-Levi [2015], but our proof uses the [CDW '16] framework and is more modular.

additive setting, i.e., BREV-PPC(F) = BREV-ADDITIVE$(\hat{F})$, as the buyer receives the full boosts in both cases. It now remains to show that Mechanism SEPARATE/FREE on F is a 4-approximation to SREV$(\hat{F})$, despite the fact that the prices in the fully-boosted additive setting are each inflated by *full* boost of getting the grand bundle.

Lemma 11.

$$\text{SREV}(\hat{F}) \leq 4\,\text{SEPARATE/FREE-PPC}(F).$$

Proof. First, we derive a lower bound on the revenue from Mechanism SEPARATE/FREE for any partition of the items into free and priced. What revenue do we yield for the partition $(\mathcal{F}, \bar{\mathcal{F}})$? Recall that for every item $i \in \bar{\mathcal{F}}$, the price posted is $\eta_i(\mathcal{F}) \cdot r_i^*$. The probability that the buyer purchases item i is at least $\Pr[t_i \geq r_i^*] = 1 - F_i(r_i^*)$, because the buyer receives the boost $\eta_i(\mathcal{F})$ from all the free items with certainty. If the buyer also purchases other items, it will only increase the buyer's value for buying item i, so the probability of purchasing item i can only increase. Hence, the revenue of mechanism SEPARATE/FREE under this particular partition $(\mathcal{F}, \bar{\mathcal{F}})$ is at least

$$\sum_{i \in \bar{\mathcal{F}}} \eta_i(\mathcal{F}) \cdot r_i^* \cdot (1 - F_i(r_i^*)) = \sum_{i \in \bar{\mathcal{F}}} \eta_i(\mathcal{F}) R_i.$$

Now we construct a graph and show that the revenue of Mechanism SEPARATE/FREE under any partition $(\mathcal{F}, \bar{\mathcal{F}})$ of the items is at least the weight of a corresponding directed cut in the following graph. Consider the graph with vertices $[m]$ corresponding to the m items, where directed edge (j, i) has weight $w_{j,i} := \eta_{ij} \cdot R_i$, where R_i is the optimal revenue for selling only item i. The graph also contains a source node s, where for all items $i \in [m]$, the edge (s, i) has weight $w_{s,i} = R_i$. (This will account for the coefficient 1 for the base valuation of the item.) The weight of the directed cut from $\mathcal{F} + \{s\}$ to $\bar{\mathcal{F}}$ is precisely:

$$\sum_{i \notin \mathcal{F}} \sum_{j \in \mathcal{F} + \{s\}} w_{j,i} = \sum_{i \in \bar{\mathcal{F}}} \left(1 + \sum_{j \in \mathcal{F}} \eta_{ij}\right) R_i = \sum_{i \in \bar{\mathcal{F}}} \eta_i(\mathcal{F}) R_i.$$

Hence, for any partition of free and priced items $(\mathcal{F}, \bar{\mathcal{F}})$, the weight of the directed cut from $\mathcal{F} + \{s\}$ to $\bar{\mathcal{F}}$ gives a lower bound on the revenue yielded by Mechanism SEPARATE/FREE for this partition.

We construct our free set by placing each item independently and uniformly at random into $\mathcal{F}$ or $\bar{\mathcal{F}}$. The expected weight of the corresponding random cut from

$\mathcal{F} + \{s\}$ to $\bar{\mathcal{F}}$ is at least $\frac{1}{4}\sum_{i\in[m]}\eta_i([m])\cdot R_i = \frac{1}{4}\text{SREV}_{\eta_{[m]}\circ t}$. To see this, observe that for every pair of items (j, i), the cut gets the weight of $\eta_{ij}R_i$ from this edge whenever $j \in \mathcal{F}$ and $i \notin \mathcal{F}$, which occurs with probability $\frac{1}{4}$. The cut also gets a weight of R_i whenever $i \in \bar{\mathcal{F}}$, which happens with probability $\frac{1}{2}$.

□

Theorem 20 now follows from Lemmas 10 and 11, and Theorem 33 with $a = 1$:

$$\begin{aligned}\text{OPT-PPC}(F) &\leq \text{OPT-ADDITIVE}(\hat{F})\\ &\leq 2\,\text{SREV}(\hat{F}) + 4\,\text{BREV}(\hat{F})\\ &\leq 8\,\text{SEPARATE/FREE-PPC}(F) + 4\,\text{BREV-PPC}(F).\end{aligned}$$

5.3.2 Proof of the Benchmark

We now prove Lemma 10: that the optimal revenue from the proportional complements setting is bounded by the optimal revenue from the fully-boosted additive setting. Again, while this is intuitive, revenue non-monotonicities make it unclear how to execute a direct proof. Instead, we use the machinery from the Lagrangian duality framework of Cai et al. [2016] to give a "dual-covering" argument. While the argument is simple and easy-to-see for those familiar with the machinery, the machinery itself is not easy.

First, we formulate the (primal) optimization problem: maximize revenue subject to incentive-compatibility, individual rationality, and feasibility. We have Lagrangian dual variables, denoted by λ, corresponding to each IC constraint, i.e., corresponding to each pair of types (t, t'). Then the Lagrangian duality framework states that, via strong duality, optimal revenue is equal to the optimal dual minimization problem, and upper bounded by any feasible dual.

Of the vast array of works that use the Lagrangian duality framework to achieve an upper bound for approximation [Cai et al., 2016; Cai and Zhao, 2017; Brustle, Cai, Wu, and Zhao, 2017; Eden et al., 2017b; Eden, Feldman, Friedler, Talgam-Cohen, and Weinberg, 2017a; Fu, Liaw, Lu, and Tang, 2017; Liu and Psomas, 2017], the standard approach used by almost all of them is to select dual variables for the setting at hand that naturally split the upper bound into terms that can be bounded by a few simple mechanisms. Then, the bulk of the work remains in bounding the unique terms with the correct mechanisms. Here, however, it is not even clear how to chose a set of dual variables that induces a good upper bound due the complementarities across items. We take a different path. We first create a new proxy additive setting, where buyers' valuations are fully-boosted. We then argue that the optimal revenue in our setting is upper bounded by the optimal revenue in the boosted additive setting. As the buyers' valuations in the boosted setting "dominate" the original buyers' valuations, the claim is intuitively true. However, due to revenue non-monotonicities, this intuition does not directly translate to a proof. We rely on duality to prove the claim. We show that the optimal dual in the original setting is at most the optimal dual in the fully-boosted additive setting, which by strong duality, is equal to the optimal revenue. This step is the only place we use duality and the rest of the analysis all happen in the primal/mechanism space.

We use $\phi_i(t) := t_i - \frac{1}{f(t)} \sum_{t'}(t'_i - t_i)\lambda(t', t)$ as the "virtual value function" given by λ. Let $f(t)$ denote the probability that the type t is realized. (We assume discrete distributions for simplicity of notation.) We denote the set of feasible allocations by $\mathcal{P}$—this is just the set that allocates at most one unit of each good. The following lemma is a direct application of Theorem 4.4 of Cai and Zhao [2017] to our setting and gives the optimal revenue in terms of these dual variables.

Lemma 12.

$$\textsc{Opt-mpph} = \min_{\lambda \geq 0} \max_{x \in \mathcal{P}} \sum_i \sum_t f(t)\phi_i(t) \sum_{S:i \in S} x_S(t)\eta_i(S).$$

This lemma allows us to move back and forth between the revenue in the primal space and a bound in the dual space.

Proof. Theorem 4.4 of Cai and Zhao [2017] states that the optimal revenue from a buyer with type $t \in T$ and *any* valuation $v(t, S)$ for the set S is as follows, where $x(t, S)$ is the primal variable for the probability that the buyer receives exactly set S when he reports type t:

$$\textsc{Opt-}v(\cdot,\cdot) = \min_{\lambda \geq 0} \max_{x \in \mathcal{P}} \sum_t f(t)\Phi(t, S)x_S(t)$$

where

$$\Phi(t, S) = v(t, S) - \frac{1}{f(t)} \sum_{t' \in T} \lambda(t', t)(v(t', S) - v(t, S)).$$

In our setting, we have that $v(t,S) = \sum_{i \in S} \eta_i(S) t_i$. Thus

$$\begin{aligned}
\Phi(t,S) &= v(t,S) - \frac{1}{f(t)} \sum_{t' \in T} \lambda(t',t)(v(t',S) - v(t,S)) \\
&= \sum_{i \in S} \eta_i(S) t_i - \frac{1}{f(t)} \sum_{t' \in T} \lambda(t',t) \left(\sum_{i \in S} \eta_i(S) t'_i - \sum_{i \in S} \eta_i(S) t_i \right) \\
&= \sum_{i \in S} \eta_i(S) \left(t_i - \frac{1}{f(t)} \sum_{t' \in T} \lambda(t',t)\,(t'_i - t_i) \right) \\
&= \sum_{i \in S} \eta_i(S) \phi_i(t)
\end{aligned}$$

and the above claim holds.[6] Note that this also applies to the additive setting, where for all i, $\eta_{ij} = 0$ for all j and $\eta_i(S) = 1$. $\square$

We first relate Opt-ppc to the optimal revenue for an instance of additive valuations; in essence we just multiply the value t_i by $\eta_i([m])$. We set up some notation first. Define $\eta_{[m]}$ to be the vector whose i^{th} coordinate is $(\eta_{[m]})_i = \eta_i([m])$, and let $\eta_{[m]} \circ t$ be the Hadamard product of the vector $\eta_{[m]}$ and the vector t. Let $\hat{F}$ be the distribution where $\eta_{[m]} \circ t$ is drawn identically to t in $F = \Pi_i F_i$, i.e., $\hat{f}\left(\eta_{[m]} \circ t\right) = f(t)$. We refer to this setting as the fully-boosted additive setting.

Proof of Lemma 10. For each i and allocation rule x, by the monotonicity in (5.1), the boost from $[m]$ is larger than that from any set S, i.e., $\eta_i(S) \leq \eta_i([m])$. Thus, we have that

$$\sum_{S: i \in S} x_S(t) \eta_i(S) \leq \eta_i([m]) \sum_{S: i \in S} x_S(t) = \eta_i([m]) \pi_i(t), \tag{5.2}$$

where we define $\pi_i(t) := \sum_{S: i \in S} x_S(t)$ to be the probability that item i is allocated to a buyer of type t. We now have the following sequence of equalities and inequalities. The first line uses Lemma 12 to move to the dual space. We would like to replace $\eta_i(S)$ by $\eta_i([m])$ everywhere (using (5.2)), but this is not possible since the virtual value function may be negative on some types. Lines 2 and 3 do this by using only non-negative virtual valuations as an upper bound. We use z^+ to denote $\max\{z, 0\}$ for any real number z. In line 4 we can bring back the original (possibly negative) virtual value function because in order to maximize this quantity, the optimal π

[6] The theorem from Cai and Zhao [2017] also holds for multiple buyers, as does a restatement of Lemma 12; we only state it for a single buyer for simplicity.

must set $\pi_i(t) = 0$ when $\phi_i(t) < 0$. Line 5 then moves to the dual space for the fully-boosted additive setting, by suitably defining the dual variables there. (The exact duals are defined below.) Line 6 uses Lemma 12 once again to come back to the primal, $\textsc{Opt-additive}(\hat{F})$.

$$
\begin{aligned}
\textsc{Opt-ppc}(F) &= \min_{\lambda \geq 0} \max_{x \in \mathcal{P}} \sum_i \sum_t f(t)\phi_i(t) \sum_{S: i \in S} x_S(t)\eta_i(S) && \text{by Lemma 12} \\
&\leq \min_{\lambda \geq 0} \max_{x \in \mathcal{P}} \sum_i \sum_t f(t) \left(\phi_i(t)\right)^+ \sum_{S: i \in S} x_S(t)\eta_i(S) \\
&\leq \min_{\lambda \geq 0} \max_{\pi} \sum_i \sum_t f(t) \left(\phi_i(t)\right)^+ \cdot \eta_i([m])\pi_i(t) && \text{by (5.2)} \\
&= \min_{\lambda \geq 0} \max_{\pi} \sum_i \sum_t f(t)\phi_i(t)\eta_i([m])\pi_i(t) \\
&= \min_{\lambda \geq 0} \max_{\pi} \sum_i \sum_{\eta_{[m]} \circ t} \hat{f}\left(\eta_{[m]} \circ t\right) \hat{\phi}_i\left(\eta_{[m]} \circ t\right) \pi_i\left(\eta_{[m]} \circ t\right) && \text{by (5.3)} \\
&= \textsc{Opt-additive}(\hat{F}) && \text{by Lemma 12.}
\end{aligned}
$$

The equality in line 5 is true because if we set the dual variable $\hat{\lambda}(\eta_{[m]} \circ t', \eta_{[m]} \circ t) = \lambda(t', t)$ in the fully-boosted additive setting, $\hat{\lambda}$ still corresponds to a feasible dual variable[7]. Therefore, it induces the following virtual value function:

$$
\begin{aligned}
\hat{\phi}_i\left(\eta_{[m]} \circ t\right) &= \eta_i([m]) \circ t_i - \frac{1}{\hat{f}\left(\eta_{[m]} \circ t\right)} \sum_{\eta_{[m]} \circ t'} \left(\eta_i([m])t_i' - \eta_i([m])t_i\right) \hat{\lambda}\left(\eta_{[m]} \circ t', \eta_{[m]} \circ t\right) \\
&= \eta_i([m])t_i - \frac{1}{f(t)} \sum_{t'} \eta_i([m])\left(t_i' - t_i\right) \lambda(t', t) \\
&= \eta_i([m])\phi_i(t). && (5.3)
\end{aligned}
$$

□

5.3.3 XOS Complementarities

For simplicity, our analysis is written for additive boosts. However, the extension to XOS boosts is fairly straight-forward. As shown in (5.1), XOS boosts are also monotone, so the upper bound from using $\eta_i([m])$ holds. We modify our graph construction from the proof of Lemma 11 as follows. Define $\ell_i^* \in \text{argmax}_{\ell \in [K]} \sum_{j \in [m] \setminus \{i\}} \eta_{ij}^{\ell}$;

[7] For readers familiar with CDW '16, $\hat{\lambda}$ still corresponds to a flow.

then $\eta_i([m]) = 1 + \sum_{j\in[m]\setminus\{i\}} \eta_{ij}^{\ell_i^*}$. Then in the XOS analysis, the directed edge (j,i) has weight $w_{j,i} := \eta_{ij}^{\ell_i^*} \cdot R_i$. A cut from $\{s\} \cup \mathcal{F}$ to $\bar{\mathcal{F}}$ will have thus have weight

$$\sum_{i\notin\mathcal{F}} \sum_{j\in\mathcal{F}+\{s\}} w_{j,i} = \sum_{i\in\bar{\mathcal{F}}} \left(1 + \sum_{j\in\mathcal{F}} \eta_{ij}^{\ell_i^*}\right) R_i \le \sum_{i\in\bar{\mathcal{F}}} \eta_i(\mathcal{F}) R_i.$$

That is, the weight of the cut is a lower bound on the revenue of the mechanism with free set $\mathcal{F}$ and items in $\bar{\mathcal{F}}$ priced accordingly, using the actual $\eta_i(\mathcal{F})$'s. Since a uniformly random $\mathcal{F}$ guarantees a cut of weight $\frac{1}{4}\sum_i \eta_i([m]) \cdot R_i$ in expectation, then the expected revenue is again at least as high.

Similarly, in Lemmas 14 and 15, the same modification of using $w_{T,i} := \eta_{iT}^{\ell_i^*}$ on edges (T,i) will guarantee that the weight of any cut is again a lower bound on the corresponding SEPARATE/FREE revenue, so our random cut constructions give the same guarantees under XOS boosts as well.

Finally, it is not hard to see that even when the boosts are XOS functions, the revenue of selling the grand bundle is still the same as the fully-boosted additive BREV$(\hat{F})$.

Theorem 21. *The better of selling the grand bundle and Mechanism* SEPARATE/FREE *is a 12-approximation to the optimal revenue for XOS complementarities.*

5.4 Extension to MPPH

In this section, we show how to extend the mechanism and the analysis to the more general proportional positive hypergraphic (PPH) valuation class. Recall that η_{iT} may be defined for any subset $T \in [m] \setminus \{i\}$, and that $\eta_i(S) = 1 + \sum_{T\subseteq S\setminus\{i\}} \eta_{iT}$. Also recall that k is the directed-positive-rank of the hypergraph, and d is the maximum-out-degree. The extensions to the boosts being a maximum over many hypergraphs (MPPH) is covered in our analysis in Section 5.3.3.

The distribution for the fully-boosted additive setting $\hat{F}$ is defined as before, except with $\eta_i([m])$ defined according to the PPH valuations.

Lemma 13 states that the fully-boosted additive setting is again a crude upper bound on revenue; it is the analog of Lemma 10 for PPH valuations and can be proven similarly.

Lemma 13.

$$\textsc{Opt-pph}(F) \leq \textsc{Opt-additive}(\hat{F}).$$

Next, we prove an analog of Lemma 11 which shows that we can obtain a $4k$-approximation to $\textsc{srev}(\hat{F})$.

Lemma 14.

$$\textsc{srev}(\hat{F}) \leq 4k\, \textsc{separate/free-pph}(F).$$

Proof. We use a random construction of the free set, and we show that the expected revenue of our mechanism is at least a $1/4k$-fraction of $\textsc{srev}(\hat{F})$. Each item independently is free (in $\mathcal{F}$) with probability $(1-\frac{1}{2k})$, and otherwise it is priced. By definition of the directed-positive-rank, for every given η_{iT}, $|T| \leq k$. Then for any such T, all items in T appear simultaneously in $\mathcal{F}$ with probability $(1-\frac{1}{2k})^{|T|} \geq (1-\frac{1/2}{k})^k \geq \frac{1}{2}$. In addition, every item $i \in \bar{\mathcal{F}}$ with probability $\frac{1}{2k}$.

Consider the graph construction where a directed edge (T, i) has weight $w_{T,i} := \eta_{iT} \cdot R_i$ and we have an edge (s, i) for every item i with weight $w_{s,i} := R_i$. Every edge is cut from $\{s\} \cup \mathcal{F}$ to $\bar{\mathcal{F}}$ with probability $\geq \frac{1}{2} \cdot \frac{1}{2k} = \frac{1}{4k}$. The expected weight of the cut from $\{s\} \cup \mathcal{F}$ to $\bar{\mathcal{F}}$ is then $\geq \frac{1}{4k} \sum_i \eta_i([m]) \cdot R_i = \frac{1}{4k} \cdot \textsc{srev}(\hat{F})$.

Again, as in the proof of Lemma 11, we observe that the expected revenue of Mechanism $\textsc{separate/free}$ with partition $(\mathcal{F}, \bar{\mathcal{F}})$ achieves at least as much revenue as the directed cut from $\{s\} + \mathcal{F}$ to $\bar{\mathcal{F}}$, and thus the mechanism obtains the $\frac{1}{4k}$-approximation. □

We prove in the next Lemma that there is a different way to choose the free set to obtain a $4d$-approximation to $\textsc{srev}(\hat{F})$.

Lemma 15.

$$\textsc{srev}(\hat{F}) \leq 4d\, \textsc{separate/free-pph}(F).$$

Proof. When the hypergraph has maximum-out-degree d, that is, d is the largest number of edges directed out of any item, a slightly different random construction of the free set gives a $4d$-approximation to $\textsc{srev}(\hat{F})$. For each hyperedge (T, i), with probability $\frac{1}{2d}$, we place *all* items $j \in T$ into the free set. We run this process for every hyperedge (in some arbitrary order). If, after this process, an item j is not assigned to the free set, then item j is priced (placed into $\bar{\mathcal{F}}$). For any item i, the item is priced when none of the (at most d) edges that are directed from a set which contains i are placed into the free set, which occurs with probability at least $(1 - \frac{1}{2d})^d \geq \frac{1}{2}$. The probability of i being free is of course at least $\frac{1}{2d}$.

Then any edge (T, i) crosses the cut from $\{s\} + \mathcal{F}$ to $\bar{\mathcal{F}}$ with probability at least $\frac{1}{2d} \cdot \frac{1}{2} = \frac{1}{4d}$. Then by the same analysis as in the proof of Lemma 14, the expected weight of the cut from $\{s\} + \mathcal{F}$ to $\bar{\mathcal{F}}$ is at least $\frac{1}{4d} \sum_i \eta_i([m]) \cdot R_i = \frac{1}{4d} \cdot \textsc{srev}(\hat{F})$, which is again a lower bound on the expected revenue of Mechanism SEPARATE/FREE with partition $(\mathcal{F}, \bar{\mathcal{F}})$. □

Together, this gives

$$\begin{aligned} \textsc{Opt-pph} \leq \textsc{Opt}(\hat{F}) &\leq 2\,\textsc{srev}(\hat{F}) + 4\,\textsc{brev}(\hat{F}) && \text{Lemma 13 and Theorem 33} \\ &\leq 8\min\{k, d\}\,\textsc{separate/free-pph} + 4\,\textsc{brev-pph}. && \text{Lemmas 14 and 15} \end{aligned}$$

Theorem 22. *The better of selling the grand bundle and Mechanism* SEPARATE/FREE *for* PPH *valuations, with directed-positive-rank k and maximum-out-degree d, is an $(8\min\{d, k\}+4)$-approximation to the optimal revenue.*

The analysis in Section 5.3.3 generalizes the guarantees to MPPH (from additive to XOS boosts).

5.4.1 Lower Bound of $O\left(\frac{1}{k}\mathrm{OPT}(\widehat{F})\right)$

In our analysis, we make two relaxations. First, we relax our benchmark from OPT-MPPH to the upper bound of $\textsc{Opt}(\hat{F})$. Second, we lower bound the revenue of our SEPARATE/FREE mechanism by undercounting the probabilities of sale.

It is extremely difficult to reason about the probability that a buyer will be interested in buying an item (or a set of items): her value may only be high enough if she buys multiple bundles simultaneously, or she may purchase a bundle even though her value for it is low because it improves her value for other bundles. Instead, we undercount this probability in the following manner: when the buyer is deciding whether to take a priced bundle of items B, we suppose that she only counts the boosts between items within that bundle B and the boost from the free items in $\mathcal{F}$. We refer to this lower bound on revenue as the *proxy revenue*.

In this section, we show that with respect to these two relaxations, for a reasonable class of simple mechanisms which includes ours, there exists an instance such that the proxy revenue of every mechanism from the class is a factor of k off from $\textsc{Opt}(\hat{F})$.

Note that this does not imply that the proxy revenue of these mechanisms is far from Opt-mpph, as we do not know how far Opt-mpph is from the benchmark of $\text{Opt}(\hat{F})$; we also do not know how far the proxy revenue may be from the actual revenue.

Definition 14. A mechanism is from the class of *Bundle Pricing Mechanisms $\mathcal{B}$* if it computes prices as follows. The mechanism determines a partition of items into y priced bundles of size $n_1, \ldots, n_y$ and one free set $\mathcal{F}$. The j^{th} bundle B_j is priced at its monopoly reserve when counting (1) the boosts of the complementarities within the bundle and (2) the boosts to the items in B_j from the free set $\mathcal{F}$.

Theorem 23. *Among Bundle Pricing Mechanisms $\mathcal{B}$, no mechanism has proxy revenue better than $O\left(\frac{1}{k}\,\text{Opt}(\hat{F})\right)$, and* separate/free *with a random free set $\mathcal{F}$ achieves this.*

Proof. Consider the following instance. There are m items, and the buyer's type for item $i \in [m]$ is

$$t_i = \begin{cases} 2^i & \text{w.p. } 2^{-i} \\ 0 & \text{otherwise.} \end{cases}$$

For every size-k set $T \in \binom{[m]}{[k]}$, for all items $i \notin T$, we have that $\eta_{iT} = c := \frac{m}{2\binom{m-1}{k}}$. That is, the market structure is the directed complete graph of hyperedges of size exactly k. Any other hyperedge (T, i) where $|T| \neq k$ has weight $\eta_{iT} = 0$. In total, there are $\binom{m-1}{k}$ edges of weight $\frac{m}{2\binom{m-1}{k}}$ into each item i, thus $\eta_i([m]) = 1 + \frac{m}{2}$.

Under these valuations and market parameters, for the random free set construction described in the previous section (pricing any item with probability $\frac{1}{2k}$), we get proxy revenue at least $(1 + \frac{m}{2})m\frac{1}{4k} = \frac{m+m^2/2}{4k}$.

We now show that the proxy revenue of every mechanism from $\mathcal{B}$ is $O(\frac{m^2}{k})$, and is thus no better than a constant factor times the proxy revenue of our mechanism.

Lemma 16. *In the above construction, for every bundle of n items and a free set $\mathcal{F}$ of size $|\mathcal{F}| = w$ items, the proxy revenue of the bundle is $O\left(\left(\binom{n}{k} + \binom{w}{k}\right) \cdot c\right)$, where $c = \frac{m}{2\binom{m-1}{k}}$.*

Proof. We count the boosts that are incorporated into the proxy revenue for any item: from within its bundle, and from the free set. First, for any item i within some bundle B of size n, the boosts from within the bundle are exactly $\sum_{T \subseteq B \setminus \{i\}:|T|=k} \eta_{iT} = \binom{n-1}{k} \cdot c$. Then, the boosts that i gets from the free set are $\sum_{T \subseteq \mathcal{F}:|T|=k} \eta_{iT} = \binom{w}{k} \cdot c$. Together,

i's boosts accounted for in the proxy revenue are

$$\eta(B+\mathcal{F}) := \quad \eta_i(B+\mathcal{F}) = \left(\binom{n-1}{k} + \binom{w}{k}\right) \cdot c.$$

We now show that for any bundle B of n items, the proxy revenue is at most $4 \cdot \eta(B+\mathcal{F})$. According to the way we undercount probability for the proxy revenue, for any ℓ, the (undercounted) probability that the buyer's value for bundle B is greater than $\eta(B+\mathcal{F})2^\ell$ is at most the probability that he has value at least 2^ℓ in base valuations, which is $\sum_{i=\ell}^{m} 2^{-i} \leq 1/2^{\ell-1}$.

Therefore, for any price for this bundle $p_B \in [\eta(B+\mathcal{F})2^\ell, \eta(B+\mathcal{F})2^{(\ell+1)}]$, the expected proxy revenue for this bundle is no more than $\eta(B+\mathcal{F})2^{(\ell+1)} \cdot 2^{-(\ell-1)} = 4 \cdot \eta(B+\mathcal{F})$. □

1. The proxy revenue for selling separately is m. This is the proxy revenue earned from optimally selling the m items separately, without giving any item out for free. Posting a price of 2^i for each item i earns expected proxy revenue 1 for each of the m items.

2. BREV $\leq 2m$, by the proof of Lemma 16 when $n = m$ and $w = 0$.

3. For any mechanism $\mathcal{M} \in \mathcal{B}$, the proxy revenue of $\mathcal{M}$-PPH is no more than $O(m^2/k)$. Consider the mechanism that offers a free set of size w to the buyer, and then splits the remaining $m - w$ items into y bundles where the j^{th} bundle is of size n_j. According to Lemma 16, the mechanism's proxy revenue is

$$O\left(\left(\binom{w}{k} \cdot y + \sum_{j=1}^{y} \binom{n_j}{k}\right) \cdot c\right).$$

Clearly, $\sum_{j=1}^{y} \binom{n_j}{k} \leq \binom{\sum_{j=1}^{y} n_j}{k} = \binom{m-w}{k}$. By definition of c, $O\left(\binom{m-w}{k} \cdot c\right) = O(m)$.

Next, we bound $\binom{w}{k} \cdot y$ by $\frac{w^k}{k!} \cdot (m-w)$. Then, by the AM-GM inequality,

$$w^k \cdot (m-w) \quad = \quad \left(\frac{w}{k}\right)^k \cdot (m-w) \cdot k^k \quad \leq \quad \left(\frac{m}{k+1}\right)^{k+1} \cdot k^k \quad = \quad O\left(\frac{m^{(k+1)}}{k+1}\right).$$

Combining everything, we have that

$$c \cdot \binom{w}{k} \cdot y \leq O\left(\frac{c}{k!} \cdot \frac{m^{(k+1)}}{k+1}\right) = O\left(\frac{m^2}{k}\right),$$

where again the definition of c kills the factor of $\frac{m^k}{k!}$.

□

5.5 A common generalization

As observed earlier, our model captures scenarios where the additional value from a combination of items depends on the base values for the items, whereas the common PH model captures scenarios where this is independent. We now present a common generalization of these two models. Consider the hypergraphic representation of a valuation function, i.e., where the valuation function is represented by

$$v(S) = \sum_{T \subseteq S} v_T$$

for some values v_T; the T for which $v_T > 0$ are the hyperedges of the underlying hypergraph. Our model can be thought of as a special case where v_T is a linear combination of the base values for the items in T :

$$v_T = \sum_{i \in T} \eta_{iT} v_i.$$

More generally, one could have an arbitrary linear transformation from the type space to the hypergraphic representation: let $t = (t_1, t_2, \ldots, t_d)$ be the type, for some dimension d, and

$$v_T = \sum_{i \in [d]} \eta_{iT} t_i.$$

An interpretation of this model is that, for each $i \in [d]$, t_i represents the buyer's value for some activity, and η_{iT} is the additional boost for that activity made possible by the buyer owning the combination of items in T. Assume that each t_i is independent of the others. This generalizes the PH model with independent v_Ts: each hyperedge corresponds to a different activity, and is boosted only by itself. The model can be further extended to XOS boosts, i.e., a maximum of many linear combinations (as in MPH). We now give an example where such a model is useful.

Example 3. Consider a computing device such as a tablet, which has multiple uses, such as browsing the web, and taking notes. A buyer's valuation for such a device can be modeled as a linear combination of his value for each of the activities it enables. Now consider an accessory such as a stylus. This makes some of the activities faster, such as taking notes. The additional value it provides can be modeled as a linear combination of values for the corresponding activities. Similarly a note-taking app also makes the note-taking activity more valuable. Moreover, it could be that a combination of a stylus and a compatible app has further added boost to the valuation for that activity.

This perspective is similar in spirit to the 'subadditive with independent items' model of Rubinstein and Weinberg [2015]. The types $(t_1, t_2, \ldots, t_m)$ are drawn from a product distribution of m spaces, one for each item; the space corresponding to each item itself can be multi-dimensional. The valuation function for a set $v(S)$ can be an arbitrary function that depends only on t_i for $i \in S$, subject to subadditivity.

What is the point of a model even more general than PH when we have seemingly strong lower bounds for PH? These lower bounds are for $\max\{\text{SREV}, \text{BREV}\}$, which are (by now) the standard pricing mechanisms for which upper bounds have been shown. While it makes sense to consider the simplest of the pricing schemes when it comes to upper bounds, lower bounds against such utterly simple pricing schemes are much less compelling. When it comes to items that are complements, where such pricing schemes may not be the most natural, such lower bounds are more of an indication that we need to study alternate pricing schemes, rather than a sign of hopelessness. A take away from our results is that suitably simple pricing schemes *could* give constant factor approximations for reasonably general valuation models with complements. It is too early to discard the hope for such results for PH and other generalizations.

Part III

Interdependent Valuations

6 Combinatorial Auctions with Interdependent Valuations

6.1 Introduction

Maximizing social welfare with private valuations is a solved problem. The classical Vickrey-Clarke-Grove (VCG) family of mechanisms [Vickrey, 1961; Clarke, 1971; Groves, 1973], of which the Vickrey second-price auction is a special case, are dominant strategy incentive-compatible and guarantee optimal social welfare in general social choice settings.

In this chapter, we consider combinatorial auctions, where each agent has a value for every subset of items, and the goal is to maximize the social welfare, namely the sum of agent valuations for their assigned bundles. As a special case of general social choice settings, the VCG mechanism solves this problem optimally, *as long as the values are independent*.

There are many settings, however, in which the independence of values is not realistic. If the item being sold has money-making potential or is likely to be resold, the values different agents have may be correlated, or perhaps even common. A classic example is an auction for the right to drill for oil in a certain location [Wilson, 1969]. Importantly, in such settings, agents may have different information about what that value actually is. For example, the value of an oil lease depends on how much oil there actually is, and the different agents may have access to different assessments about this. Consequently, an agent might change her own estimate of the value of the oil lease given access to the information another agent has. Similarly, if an agent had access to the results of a house inspection performed by a different agent, that might change her own estimate of the value of a house that is for sale.

The following model due to Milgrom and Weber [1982], described here for single-item auctions, has become standard for auction design in such settings. These are known as *interdependent value settings* (IDV) [1] and are defined as follows:

This chapter is based on joint work with Alon Eden, Michal Feldman, Amos Fiat, and Anna Karlin in a paper titled "Combinatorial Auctions with Interdependent Valuations: SOS to the Rescue" which appeared at EC 2019 [EFFGK '19] and was awarded Best Paper with a Student Lead Author.

[1] See also [Krishna, 2009; Milgrom, 2004].

- Each agent i has a real-valued, private *signal* s_i. The set of signals $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ may be drawn from a (possibly) correlated distribution.

  The signals summarize the information available to the agents about the item. For example, when the item to be sold is a house, the signal could capture the results of an inspection and privately collected information about the school district. In the setting of oil drilling rights, the signals could be information that each companies' engineers have about the site based on geologic surveys, etc.

- The *value of the item* to agent i is a function $v_i(\mathbf{s})$ of the signals (or information) of *all* agents.

  A typical example is when $v_i(\mathbf{s}) = s_i + \beta \sum_{j \neq i} s_j$, for some $\beta \leq 1$. This type of valuation function captures settings where an agent's value depends both on how much he likes the item (s_i) and on the resale value which is naturally estimated in terms of how much other agents like the item ($\sum_{j \neq i} s_j$) [Myerson, 1981].

In the economics literature, interdependent settings have been studied for about 50 years (with far too many papers to list; for an overview, see [Krishna, 2009]). Within the theoretical computer science community, interdependent (and correlated) settings have received less attention (see Section 6.1.4 for further discussion and references).

6.1.1 Maximizing Social Welfare

Consider the goal of maximizing social welfare in interdependent settings. Here, a direct revelation mechanism consists of each agent i reporting a bid for their private signal s_i, and the auctioneer determining the allocation and payments. (It is assumed that the auctioneer knows the form of the valuation functions $v_i(\cdot)$.)

In interdependent settings, it is not possible[2] to design dominant-strategy incentive-compatible auctions, since an agent's value depends on *all* of the signals, so if, say, agent i misreports his signal, then agent j might win at a price above her value if she reports truthfully. The next strongest equilibrium notion one could hope for is

[2]Except perhaps in degenerate situations.

to maximize efficiency in ex-post equilibrium: bidding truthfully is an *ex-post equilibrium* if an agent does not regret having bid truthfully, given that other agents bid truthfully. In other words, bidding truthfully is a Nash equilibrium for every signal profile.[3] A strong impossibility result due to Jehiel and Moldovanu [2001] shows that with *multi-dimensional* signals, maximizing welfare is generically impossible even in Bayes-Nash equilibrium.[4]

For single-item auctions with single-dimensional signals, a characterization of ex-post incentive compatibility in the IDV setting is known, analogous to Myerson's characterization for the independent private values model (e.g., Roughgarden and Talgam-Cohen [2016]). The characterization says that there are payments that yield an ex-post incentive-compatible mechanism if and only if the corresponding allocation rule is monotone in each agent's signal, when all other signals are held fixed. Maximizing efficiency in ex-post equilibrium is also provably impossible unless the valuation functions $v_i(\mathrm{s})$ satisfy a technical condition known as the *single-crossing condition* [Milgrom and Weber, 1982; d'Aspremont and Gérard-Varet, 1982; Maskin, 1992; Ausubel et al., 1999; Dasgupta and Maskin, 2000; Athey, 2001; Bergemann, Shi, and Välimäki, 2009; Chawla, Fu, and Karlin, 2014; Che, Kim, and Kojima, 2015; Li, 2016; Roughgarden and Talgam-Cohen, 2016]. I.e., the influence of agent i's signal on his own value is at least as high as its influence on other agents' values, when all other signals s_{-i} are held fixed [5]. When the single-crossing condition holds, there is a generalization of VCG that maximizes efficiency in ex-post equilibrium. (See [Crémer and McLean, 1985, 1988; Krishna, 2009].)

Unfortunately, the single crossing condition does not generally suffice to obtain optimal social welfare in settings beyond that of a single item auction with single-dimensional signals. It is insufficient in fairly simple settings, such as two-item, two-bidder auctions with unit-demand valuations (see Section C.1), or single-parameter settings with downward-closed feasibility constraints (see Section C.2).

Moreover, there are many relevant single-item settings where the single-crossing

[3] Note that, of course, every ex-post equilibrium is a Bayes-Nash incentive compatible equilibrium, but not necessarily vice versa, and therefore ex-post equilibria are much more robust: they do not depend on knowledge of the priors and bidders need not think about how other bidders might be bidding. This increases our confidence that an ex-post equilibrium is likely to be reached.

[4] For more details on this and other related work, see Section 6.1.4.

[5] This implies that given signals s_{-i}, if agent i has the highest value when $s_i = s^*$, then agent i continues to have the highest value for $s_i > s^*$. This is precisely the monotonicity needed for ex-post incentive compatibility.

condition does not hold. For example, suppose that the signals indicate demand for a product being auctioned, agents represent firms, and one firm has a stronger signal about demand, but is in a weaker position to take advantage of that demand. A setting like this could yield valuations that do not satisfy the single crossing condition. For a concrete example, see Example 2 in Section 1.4.

6.1.2 Research Problems

This chapter addresses the following two issues related to social welfare maximization in the interdependent values model:

1. To what extent can the optimal social welfare be approximated in interdependent settings that do not satisfy the single-crossing condition?
2. How far beyond the single item, single-dimensional setting can we go?

   Given the impossibility result of Jehiel and Moldovanu [2001], we ask if it is possible to *approximately* maximize social welfare in *combinatorial auctions with interdependent values*?

The first question was recently considered by Eden et al. [2018] who gave two examples pointing out the difficulty of approximating social welfare without single crossing. Example 4 shows that even with two bidders and one signal, there are valuation functions for which no deterministic auction can achieve *any* bounded approximation ratio to optimal social welfare.

Example 4 (No bound for deterministic auctions Eden et al. [2018])**.** A single item is for sale. There are two players, A and B, only A has a signal $s_A \in \{0, 1\}$. The valuations are

$$\begin{aligned} v_A(0) &= 1 \qquad v_B(0) = 0 \\ v_A(1) &= 2 \qquad v_B(1) = H, \end{aligned}$$

where H is an arbitrary large number. If A doesn't win when $s_A = 0$, then the approximation ratio is infinite. On the other hand, if A does win when $s_A = 0$, then by monotonicity, A must also win at $s_A = 1$, yielding a $2/H$ fraction of the optimal social welfare.

The next example can be used to show that there are valuation functions for which no randomized auction performs better (in the worst case) than allocating to a random bidder (*i*.e., a factor n approximation to social welfare), even if a prior over the signals is known.

Example 5 (n lower bound for randomized auctions Eden et al. [2018])**.** There are n bidders $1, \ldots, n$ that compete over a single item. For every agent i, $s_i \in \{0, 1\}$, and

$$v_i(\mathbf{s}) = \prod_{j \neq i} s_j + \epsilon \cdot s_i \quad \text{for } \epsilon \to 0;$$

that is, agent i's value is high if and only if all other agents' signals are high simultaneously. When all signals are 1, then in any feasible allocation, there must be an agent i which is allocated with probability of at most $1/n$. By monotonicity, this means that the probability this agent is allocated when the signal profile is $\mathbf{s}' = (\mathbf{1}_{-i}, 0_i)$ is at most $1/n$ as well. Therefore, the achieved welfare at signal profile $\mathbf{s}'$ is at most $1/n + (n-1) \cdot \epsilon$, while the optimal welfare is 1, giving a factor n gap [6].

Therefore, *some* assumption is needed if we are to get good approximations to social welfare. The approach taken by Eden et al. [2018] was to define a relaxed notion of single-crossing that they called c-single crossing and then provide mechanisms that approximately maximize social welfare, where the approximation ratio depends on c and n, the number of agents.

In this chapter, we go in a different direction, starting with the observation that in Example 5, the valuations treat the signals as highly-complementary–one has a value bounded away from zero only if all other agent's signals are high simultaneously. This suggests that the case where the valuations treat the signals more like "substitutes" might be easier to handle.

We capture this by focusing on *s*ubmodular over signals (SOS) valuations. This means that for every i and j, when signals $\mathbf{s}_{-j}$ are lower, the sensitivity of the valuation $v_i(\mathbf{s})$ to changes in s_j is higher. Formally, we assume that for all j, for any s_j, $\delta \geq 0$, and for any $\mathbf{s}_{-j}$ and $\mathbf{s}'_{-j}$ such that component-wise $\mathbf{s}_{-j} \leq \mathbf{s}'_{-j}$, it holds that

$$v_i(s_j + \delta, \mathbf{s}_{-j}) - v_i(s_j, \mathbf{s}_{-j}) \geq v_i(s_j + \delta, \mathbf{s}'_{-j}) - v_i(s_j, \mathbf{s}'_{-j}).$$

[6] Eden et al. [2018] show that there exists a prior for which the n gap still holds, *even* if the mechanism knows the prior.

Many valuations considered in the literature on interdependent valuations are SOS (though this term is not used) Milgrom and Weber [1982]; Dasgupta and Maskin [2000]; Klemperer [1998]. The simplest (yet still rich) class of SOS valuations are *fully separable* valuation functions [7], where there are *arbitrary* (weakly increasing) functions $g_{ij}(s_j)$ for each pair of bidders i and j such that

$$v_i(\mathbf{s}) = \sum_{j=1}^{n} g_{ij}(s_j).$$

A more general class of SOS valuation functions are functions of the form $v_i(\mathbf{s}) = f(\sum_{j=1}^{n} g_{ij}(s_j))$, where f is a weakly increasing concave function.

We can now state the main question we study in this chapter: *to what extent can social welfare be approximated in interdependent settings with SOS valuations?* Unfortunately, Example 4 itself describes SOS valuations, so no deterministic auction can achieve any bounded approximation ratio, even for this subclass of valuations. Thus, we must turn to randomized auctions.

6.1.3 Our Results and Techniques

All of our positive results concern the design of *randomized, prior-free, universally ex-post incentive-compatible (IC), individually rational (IR) mechanisms*. Prior-free means that the rules of the mechanism makes no use of the prior distribution over the signals, thus need not have any knowledge of the prior.

Our first result provides approximation guarantees for single-parameter downward-closed settings. An important special case of this result is single-item auctions, which was the focus of Eden et al. [2018].

Theorem 26 (See Section 6.4): For every single-parameter downward-closed setting, if the valuation functions are SOS, then the `Random Sampling Vickrey` auction is a universally ex-post IC-IR mechanism that gives a 4-approximation to the optimal social welfare.

Interestingly, no deterministic mechanism can give better than an $(n-1)$-approximation for arbitrary downward-closed settings, even if the valuations are single crossing,

[7]This type of valuation function is ubiquitous in the economics literature on inderdependent settings; often with the function simply assumed to be a linear function of the signals (see, e.g., Jehiel and Moldovanu [2001]; Klemperer [1998]).

and this is tight. Recall that for a single item auction, or even multiple identical items, with single crossing valuations, the deterministic generalized Vickrey auction obtains the optimal welfare Maskin [1992]; Ausubel et al. [1999].

We then turn to multi-dimensional settings. In the most general combinatorial auction model that we consider, each agent i has a signal s_{iT} for each subset T of items, and a valuation function $v_{iT} := v_{iT}(s_{1T}, s_{2T}, \ldots, s_{nT})$. For this setting, it is not at all clear under what conditions it might be possible to maximize social welfare in ex-post equilibrium.[8]

However, rather surprisingly (see the related work section below), for the case of *separable SOS* valuations[9], we are able to extend the 4-approximation guarantee to combinatorial auctions.

Theorem 28 (See Section 6.5): For every combinatorial auction, if the valuation functions are separable-SOS, then the `Random Sampling VCG` auction is a universally ex-post IC-IR mechanism that gives a 4-approximation to the optimal social welfare.

Finally, we consider combinatorial auctions where each agent i has a single-dimensional signal s_i, but where the valuation function v_{iT} for each subset of items T is an *arbitrary* SOS valuation function $v_{iT}(s_1, \ldots, s_n)$. For this case, we show the following:

Theorems 30 and 31 (See Sections 6.6.1 and 6.6.2): Consider combinatorial auctions with single-dimensional signals, where each signal takes one of k possible values. If the valuation functions are SOS, then there exists a universally ex-post IC-IR mechanism that gives a $(k + 3)$-approximation to the optimal social welfare. If the valuations are strong-SOS [10], the approximation ratio improves to $O(\log k)$.

All of the above results, as well as our lower bounds, are summarized in Table 6.1. In addition, all of the results in this chapter generalize easily, with a corresponding degradation in the approximation ratio, to the weaker requirement of d-SOS

[8]See the related work and also Lemmas 22 and 23, which show that under one natural generalization of single-crossing to the setting of two items and two agents that are unit demand, single crossing is not sufficient for full efficiency.

[9]A valuation is separable-SOS if the valuation for an agent can be split into two parts, an SOS function of all other signals and an arbitrary function of the agents' own signal. Such valuations generalize the fully separable case discussed above. See definition 30

[10]See definition 28.

valuations [11].

6.1.3.1 Intuition for results

The fundamental tension in settings with interdependent valuations that is not present in the private values setting is the following. Consider, for example, a single item auction setting where agent 1's truthful report of her signal increases agent 2's *value*. Since, this increases the chance that agent 2 wins and may decrease agent 1's chance of winning, it might motivate agent 1 to strategize and misreport.

Our approach is to simply *prevent* this interaction. Without looking at the signals, our mechanism randomly divides the agents into two sets[12]: potential winners and certain losers. Losers never receive any allocation. When estimating the value of a potentially winning agent i, we use only the signals of losers and i's own signal(s). Thus, potential winners can not impact the estimated values and hence allocations of other potential winners. This resolves the truthfulness issue. The remaining question is: can we get sufficiently accurate estimates of the agents' values when we ignore so many signals?

The key lemma (**Lemma 18** Section 6.3) shows that we can do so, when the valuations are SOS. Specifically, for any agent i, if all agents other than i are split into two random sets A (losers) and B (potential winners), and the signals of agents in the random subset B are "zeroed out", then the expected value agent i has for the item is at least half of her true valuation. That is,

$$E_A[v_i(s_i, \mathbf{s}_A, \mathbf{0}_B)] \geq \frac{1}{2} v_i(\mathbf{s}).$$

Dealing with combinatorial settings is more involved as the truthfullness characterization is less obvious, but the key ideas of random partitioning and using the signals of certain losers remain at the core of our results.

6.1.3.2 Additional remarks

While this chapter deals entirely with welfare maximization, our results have significance for the objective of maximizing the seller's revenue. Eden et al. [2018]

[11] A valuation function is d-SOS if for all j, for all $\delta > 0$, and for any $\mathbf{s}_{-j}$ and $\mathbf{s}'_{-j}$ such that component-wise $\mathbf{s}_{-j} \leq \mathbf{s}'_{-j}$, it holds that $d \cdot (v_i(s_j + \delta, \mathbf{s}_{-j}) - v_i(s_j, \mathbf{s}_{-j})) \geq v_i(s_j + \delta, \mathbf{s}'_{-j}) - v_i(s_j, \mathbf{s}'_{-j})$.

[12] as in [Goldberg, Hartline, and Wright, 2001].

Setting	Approximation Guarantees	
Single Parameter SOS valuations Downward Closed Feasibility Single-Dimensional Signals	$\geq 1/4$ $\forall$mech. $\leq 1/2$	(Section 6.4)
Arbitrary Combinatorial SOS valuations Single-Dimensional Signals, k-sized Signal Space	$\geq 1/(k+3)$ $\forall$mech. $\leq 1/2$	(Section 6.6.1)
Arbitrary Combinatorial, Strong-SOS Valuations Single-Dimensional Signals, k-sized Signal Space	$\geq 1/(\log(k)+2)$ $\forall$mech. $\leq 1/2$	(Section 6.6.2)
Combinatorial, Separable-SOS Valuations Multi-Dimensional Signals	$\geq 1/4$ $\forall$mech. $\leq 1/2$	(Section 6.5)

Table 6.1: The table shows the approximation factors achievable for social welfare maximization with SOS and strong-SOS valuations. Similar results hold for d-approximate SOS/Strong-SOS valuations, while losing a factor that depends on d. All positive results are obtained with universally ex-post IC-IR randomized mechanisms.

give a reduction from revenue maximization to welfare maximization in single-item auctions with SOS valuations. Thus, the constant factor approximation mechanism presented in this chapter implies a constant factor approximation to the optimal revenue in single-item auctions with SOS valuations. We note that this is the first revenue approximation result that does not assume any single-crossing type assumption ([Chawla et al., 2014; Eden et al., 2018; Roughgarden and Talgam-Cohen, 2016; Li, 2016] require single crossing or approximate single crossing).

Finally, one can easily verify that, based on Yao's min-max theorem, the existence of a *randomized prior-free mechanism* that gives some approximation guarantee (in expectation over the coin flips of the mechanism) implies the existence of a *deterministic prior-dependent* mechanisms that gives the same approximation guarantee (in expectation over the signal profiles).

6.1.4 Extended Related Work

As discussed above, in single-parameter settings, there is an extensive literature on mechanism design with interdependent valuations that considers social welfare maximization, revenue maximization and other objectives. However, the vast majority of this literature assumes some kind of single-crossing condition and, in the context of social welfare, focuses on exact optimization.

There are two papers that we are aware of that study the question of how well optimal social welfare can be approximated in ex-post equilibrium without single-crossing. The first is the aforementioned paper [Eden et al., 2018] on single item auctions with interdependent valuations. They defined a parameterized version of single-crossing, termed c-single crossing, where $c > 1$ is a parameter that indicates how close is the valuation profile to satisfy single-crossing. For c-single crossing valuations, they provide a number of results including a lower bound of c on the approximation ratio achievable by any mechanism, a matching upper bound for binary signal spaces, and mechanisms that achieve approximation ratios of $(n-1)c$ and $2c^{3/2}\sqrt{n}$ (the first is deterministic and the second is randomized).

Ito and Parkes [2006] also consider approximating social welfare in the interdependent setting. Specifically, they propose a greedy contingent-bid auction (à la [Dasgupta and Maskin, 2000]) and show that it achieves a $\sqrt{m}$ approximation to the optimal social welfare for m goods, in the special case of combinatorial auctions with single-minded bidders.

For multi-dimensional signals and settings, the landscape is sparser (and bleaker) and, to our knowledge, focuses on exact social welfare maximization. Maskin [1992] has observed that, in general, no efficient incentive-compatible single item auction exists if a buyer's valuation depends on a multi-dimensional signal.

Dasgupta and Maskin [2000] extend the VCG mechanism to the *Generalized VCG* (G-VCG) mechanism to maximize social welfare in interdependent value settings as well. When buyer signals are single-dimensional for a single-item, then G-VCG attains optimal social welfare. However, when buyers may have multiple signals (even when competing for a single item), they prove that in some cases, optimal efficiency may no longer be attainable. Instead, they consider the *second-best* benchmark of maximum social welfare subject to Bayesian incentive-compatibility constraints (also known as "constrained efficiency"), and prove that G-VCG attains this second-best welfare. They extend this result to multiple goods as well, so long as the valuations satisfy a separability condition, described below, that essentially compresses the signals to a single-dimensional statistic.

The authors assume a number of assumptions on the valuation functions, the buyers' signals, and the information that is known about them. They assume that the valuation functions $v_i(\cdot)$ are continuously differentiable, strictly monotone increasing in s_i, and that they satisfy a weak version of the single-crossing assumption. That is, buyer i and j are tied for the highest values at some signal profile $\vec{s}$, then $v_i(\cdot)$ must be more sensitive to i's signal at $\vec{s}$ than $v_j(\cdot)$. Formally,

$$\frac{\partial}{\partial s_i} v_i(s_i, \vec{s}_{-i}) > \frac{\partial}{\partial s_i} v_j(s_i, \vec{s}_{-i}).$$

The intuition for why efficiency is not always attainable with multi-dimensional signals is as follows. Suppose that bidder i has a valuation function such that multiple profiles of his own signals reduce to the same compressed signal t and give him the same utility. However, for other bidders, different profiles give them different utility. It is impossible to incentivize bidder i to report the profile that maximizes welfare for other bidders since both achieve the same utility for him.

In the multi-dimensional signal setting, the authors add the assumption that the valuation functions are *separable*, meaning that each buyer's signals $\vec{s}_i$ are compressed by every valuation function $v_j(\cdot)$ to some one-dimensional statistic t_i^j before the statistics from the various buyers interact with one other. Under this very strong assumption, they also generalize monotonicity (that if buyer i prefers some allocation, his preference only increases in t_i^i) and weak single-crossing (if i prefers some allocation that tied as welfare-maximizing with some other allocation, then the welfare of his preferred allocation increases at least as fast in t_i^i as the other allocation).

As a direct revelation mechanism, the G-VCG mechanism would operate as follows. Bidders would report their signals directly to the auctioneer, who, knowing the valuation functions, would compute the bidders' valuation functions on all of the reported signals. Assuming the reported signals are truthful, she then has the ability to determine exactly what allocation is efficient.

However, Dasgupta and Maskin prefer not to use direct revelation mechanisms, citing them as too costly since they require the designer to know the valuation functions and signal spaces. Instead, they focus on Bayes-Nash equilibria, where buyers bid *values* rather than signals, and they develop bidding strategies in response to the reported values of other buyers. Note that a buyer must bid his value without knowing his own value, so the best he can do is apply a strategy to his signal, valuation function, and the conditional information he has about the distributions of others' signals based on their reported bids. Of course, this heavily relies on full knowledge of the prior distributions of buyers' signals. The technical approach is to look for a fixed point in the bidding strategy. At first, the authors seem to assume that buyers know the valuation functions and signal spaces well enough to reverse engineer a bidder's signal from observing their reported value. However, they point out that they need only know a fixed point in the bidding strategy. What is unclear is that the buyers would actually be able to converge upon this equilibrium.

Jehiel and Moldovanu [2001] study efficient Bayesian incentive compatible (BIC) auctions in two kinds of settings: those with multi-dimensional signals and those with one-dimensional signals. They prove impossibility results in the majority of cases for multi-dimensional signals, and characterize exactly when there exist auctions that are both efficient and BIC.

Specifically, in their model, there are k possible allocation outcomes. Each buyer i has a valuation function for each outcome k that depends on signals of buyers $j = 1, \ldots, n$ that are specific to this outcome k and to i, that is, $V_k^i(s_{ki}^1, \ldots, s_{ki}^n)$. The valuation functions are linear, e.g. $V_k^i(s_{ki}^1, \ldots, s_{ki}^n) = \sum_{j=1}^n a_{ki}^j s_{ki}^j$ where the coefficients of the signals a_{ki}^j are non-negative and common knowledge to all. Thus, their valuation functions are, in one sense, a special case of our separable valuation functions. On the other hand, they are more general in that all quantities depend on the outcome k. Thus, there are allocation externalities.

First, they study the environment where a bidder might have multiple signals for the same outcome, and prove that there are instances for which there are no efficient BIC auctions. The reason for this is similar to that from the negative result in [Dasgupta and Maskin, 2000]: the single payment does not provide sufficient incentives to extract multiple signals from the agent.

The second impossibility result pertains to the case in which an agent i only has one signal per outcome k, e.g. $s_{kj}^i = s_{kj'}^i$ for all bidders j, j' (but i has a signal for every outcome k, and thus still has multi-dimensional signals). This impossibility holds even though there is only one signal to extract per outcome, unlike the previous environment. Instances exist where in order to select the efficient alternative, the mechanism should cause a bidder to be indifferent between two alternatives at many signal profiles using payments. However, this is only possible in a BIC manner if the valuation functions happen to take on a very basic form, and if they do not, then this task becomes impossible.

The authors further characterize the conditions that a mechanism must take on in order to be BIC, and show that in the case where buyers have only one-dimensional signals, this simplifies to a monotonicity-like condition that does not depend on the prior distributions of the signals.

Jehiel, Meyer-ter Vehn, Moldovanu, and Zame [2006] go on to show that the only deterministic social choice functions that are ex-post implementable in generic mechanism design frameworks with multi-dimensional signals, interdependent valuations and transferable utilities, are constant functions.

Finally, Bikhchandani [2006] considers a single item setting with multi-dimensional signals but no allocation externalities and shows that there is a generalization of single-crossing that allows some social choice rules to be implemented ex-post.

Some work in the algorithmic mechanism design community has also focused on (approximate) revenue maximization in interdependent value settings with single-dimensional signals. Roughgarden and Talgam-Cohen [2016] assumes that the prior distributions are *affiliated*, a form of positive correlation, and solve for an optimal Myersonian-like theory under this condition. Chawla et al. [2014] choose a random set of potential winners, use the G-VCG allocation, and then apply lazy conditional reserves as in the Lookahead auction to construct a mechanism that gives a constant-factor approximation to revenue for any distributions. However, they assume a concavity assumption on the valuation functions. Both papers, like almost all prior work, assume the single-crossing condition on the valuation functions.

For further analysis and discussion of implementation with interdependent valuations, see e.g., Bergemann and Morris [2005] and McLean and Postlewaite [2015].

For further literature in computer science on interdependent and correlated values, see [Ronen, 2001; Constantin, Ito, and Parkes, 2007; Constantin and Parkes, 2007; Klein, Moreno, Parkes, Plakosh, Seuken, and Wallnau, 2008; Papadimitriou and Pierrakos, 2011; Dobzinski, Fu, and Kleinberg, 2011; Babaioff, Kleinberg, and Paes Leme, 2012; Abraham, Athey, Babaioff, and Grubb, 2011; Robu, Parkes, Ito, and Jennings, 2013; Kempe, Syrgkanis, and Tardos, 2013; Che et al., 2015; Li, 2016; Chawla et al., 2014].

6.2 Model and Definitions

6.2.1 Single Parameter Settings

In Section 6.4, we will consider single-parameter settings with interdependent valuations and downward-closed feasibility constraints. In these settings, a mechanism decides which subset of agents $1, \ldots, n$ are to receive "service" (e.g., an item). The feasibility constraint is defined by a collection $\mathcal{I} \subseteq 2^{[n]}$ of subsets of agents that may feasibly be served simultaneously. We restrict attention to *downward-closed settings*, which means that any subset of a feasible set is also feasible. A simple example is a k-item auction, where $\mathcal{I}$ is the collection of all subsets of agents of size at most k.

For these settings, we use the interdependent value model of Milgrom and Weber [1982]:

Definition 15 (single-dimensional Signals, Single Parameter Valuations)**.** Each agent j has a private signal $s_j \in \mathbb{R}^+$. The value agent j gives to "receiving service" $v_j(\mathbf{s}) \in \mathbb{R}^+$, is a function of all agents' signals $\mathbf{s} = (s_1, s_2, \ldots, s_n)$. The function $v_j(\mathbf{s})$ is assumed to be weakly increasing in each coordinate and strictly increasing in s_i.

6.2.1.1 Deterministic Mechanisms

Definition 16 (Deterministic Single Parameter Mechanisms)**.** A deterministic mechanism $M = (x, p)$ in the downward closed setting is a mapping from reported signals $\mathbf{s} = (s_1, \ldots, s_n)$ to allocations $x(\mathbf{s}) = \{x_i(\mathbf{s})\}_{1 \leq i \leq n}$ and payments $p(\mathbf{s}) = \{p_i(\mathbf{s})\}_{1 \leq i \leq n}$, where $x_i(\mathbf{s}) \in \{0, 1\}$ indicates whether or not agent i receives service and $p_i(\mathbf{s})$ is the payment of agent i. It is required that the set of agents that receive service is feasible, i.e., $\{i \mid x_i(\mathbf{s}) = 1\} \in \mathcal{I}$. (The mechanism designer knows the form of the valuation functions but learns the private signals only when they are reported.)

Definition 17 (Agent utility)**.** Given a deteministic mechanism (x, p), the *utility* of agent i when her true signal is s_i, she reports s_i' and the other agents report $\mathbf{s}_{-i}$ is

$$u_i(s_i', \mathbf{s}_{-i} | s_i) = x_i(s_i', \mathbf{s}_{-i}) v_i(s_i, \mathbf{s}_{-i}) - p_i(s_i', \mathbf{s}_{-i}).$$

Agent i will report s_i' so as to maximize $u_i(s_i', \mathbf{s}_{-i} | s_i)$. We use $u_i(\mathbf{s})$ to denote the utility when she reports truthfully, i.e., $u_i(s_i, \mathbf{s}_{-i} | s_i)$.

Definition 18 (Deterministic ex-post incentive compatibility (IC))**.** A deterministic mechanism $M = (x, p)$ in the interdependent setting is *ex-post incentive compatible* (IC) if, irrespective of the true signals, and given that all other agents report their true signals, there is no advantage to an agent to report any signal other that her true signal. In other words, assuming that $\mathbf{s}_{-i}$ are the true signals of other bidders. $u_i(s_i', \mathbf{s}_{-i} | s_i)$ is maximized by reporting s_i truthfully.

Definition 19 (Deterministic ex-post individual rationality (IR))**.** A deterministic mechanism in the interdependent setting is *ex-post individually rational* (IR) if, irrespective of the true signals, and given that all other agents report their true signals, no agent gets negative utility by participating in the mechanism.

If a deterministic mechanism is both ex-post IR and ex-post IR we say that it is ex-post IC-IR.

Definition 20. A deterministic allocation rule x is monotone if for every agent i, every signal profile of all other agents s_{-i}, and every $s_i \leq s'_i$, it holds that $x_i(s_i, s_{-i}) = 1 \Rightarrow x_i(s'_i, s_{-i}) = 1$.

Proposition 24. *[Roughgarden and Talgam-Cohen, 2016] For every deterministic allocation rule x for single parameter valuations, there exist payments p such that the mechanism (x, p) is ex-post IC-IR if and only if x_i is monotone for every agent i.*

6.2.1.2 Randomized Mechanisms

Definition 21. A randomized mechanism is a probability distribution over deterministic mechanisms.

Definition 22 (Universal ex-post IC-IR)**.** A randomized mechanism is said to be universally ex-post IC-IR if all deterministic mechanisms in the support are ex-post IC-IR.

6.2.2 Combinatorial Valuations with Interdependent Signals

Sections 6.5 and 6.6 focus on combinatorial auctions, where there are n agents and m items. In these settings, a mechanism is used to decide how the items are partitioned among the agents. We consider two models for the interdependent valuations: [13]

Definition 23 (single-dimensional Signals, Combinatorial Valuations)**.** Each agent i has a signal $s_i \in \mathbb{R}^+$. The value agent i gives to subset of items $T \subseteq [m]$, which we denote by $v_{jT}(\mathbf{s})$, is a function of $\mathbf{s} = (s_1, s_2, \dots, s_n)$.

[13]For other types of signals and interdependent valuation models, see, e.g., Jehiel and Moldovanu [2001].

Definition 24 (multi-dimensional Combinatorial Signals, Combinatorial Valuations)**.** Here, each agent has a signal for each subset of items; for any agent i, we use s_{iT} to denote agent i's signal for subset of items $T \subseteq [m]$. The value agent i gives to set T is denoted by $v_{iT}(\mathbf{s}_T)$ where $\mathbf{s}_T = (s_{1T}, s_{2T}, \ldots, s_{nT}) \in \mathbb{R}^{+n}$. We use s to denote the set of all signals $\{\mathbf{s}_T\}_{T \subseteq 2^m}$.

In both cases, each $v_{iT}(\cdot)$ is assumed to be a weakly increasing function of each signal and strictly increasing in s_i (or s_{iT} respectively), and known to the mechanism designer.

We give subsequent definitions only for multi-dimensional combinatorial signals, as single-dimensional signals can be viewed as a special case of multi-dimensional signals where $s_{iT} = s_i$ for all T.

6.2.2.1 Deterministic Mechanisms

Definition 25 (Deterministic mechanisms for combinatorial settings)**.** A deteministic mechanism $M = (x, p)$ is a mapping from reported signals s to allocations $x = \{x_{iT}\}$ (where each $x_{iT} \in \{0, 1\}$) and payments $p = \{p_{iT}\}$ for all $1 \leq i \leq n$ and $T \subset \{1, \ldots, m\}$ such that:

- Agent j is allocated the set T iff $x_{jT}(\mathbf{s}) = 1$;
- For each agent j, there is at most one T for which $x_{jT}(\mathbf{s}) = 1$;
- The sets allocated to different agents do not intersect.
- The payment for agent j when her allocation is set T is $p_{jT}(\mathbf{s})$.

Definition 26 (Agent Utility)**.** The *utility* of agent i when her signals are $\mathbf{s}_i = \{s_{iT}\}_{T \subset 2^m}$, she reports $\mathbf{s}'_i$ and the other agents report $\mathbf{s}_{-i}$ is

$$u_i(\mathbf{s}'_i, \mathbf{s}_{-i} | \mathbf{s}_i) = \sum_{T \subseteq 2^m} x_{iT}(\mathbf{s}'_i, \mathbf{s}_{-i})[v_{iT}(\mathbf{s}_{iT}, \mathbf{s}_{-iT}) - p_{iT}(\mathbf{s}'_i, \mathbf{s}_{-i})].$$

Given a mechanism $M = (x, p)$, agent i will report $\mathbf{s}'_i$ so as to maximize $u_i(\mathbf{s}'_i, \mathbf{s}_{-i} | \mathbf{s}_i)$. We use $u_i(\mathbf{s})$ to denote the utility when she reports truthfully, i.e., $u_i(\mathbf{s}_i, \mathbf{s}_{-i} | \mathbf{s}_i)$.

The definitions of *ex-post incentive compatibility* (IC) and *ex-post individually rationality* (IR) for deterministic mechanisms for combinatorial settings are the same as the appropriate definitions for single parameter mechanisms (Definitions 18 and 19 with the obvious modifications).

6.2.2.2 Randomized Mechanisms

As with single parameter mechanisms, a randomized mechanism for a combinatorial setting is a probability distribution over deterministic mechanisms for the combinatorial setting, and a randomized mechanism is said to be *universally ex post IC-IR* if all deterministic mechanisms in the support are themselves ex-post IC-IR.

6.2.3 Submodularity over signals (SOS)

As discussed in the introduction, our results will rely on an assumption about the valuation functions that we call *submodularity over signals* or SOS. The SOS (resp. strong-SOS) notion we use is the same as the weak diminishing returns (resp. strong diminishing returns) submodularity notion in [Bian, Levy, Krause, and Buhmann, 2017; Niazadeh, Roughgarden, and Wang, 2018][14]. SOS was also used in [Eden et al., 2018], generalizing a similar notion in [Chawla et al., 2014].

Definition 27 (d-approximate submodular-over-signals valuations (d-SOS valuations))**.** A valuation function $v(\mathbf{s})$ is a *d-SOS valuation* if for all j, s_j, $\delta \geq 0$,

$$\mathbf{s}_{-j} = (s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_n) \quad \text{and} \quad \mathbf{s}'_{-j} = (s'_1, \ldots, s'_{j-1}, s'_{j+1}, \ldots, s'_n)$$

such that $\mathbf{s}'_{-j}$ is smaller than or equal to $\mathbf{s}_{-j}$ coordinate-wise, it holds that

$$d \cdot \left(v(\mathbf{s}'_{-j}, s_j + \delta) - v(\mathbf{s}'_{-j}, s_j) \right) \geq v(\mathbf{s}_{-j}, s_j + \delta) - v(\mathbf{s}_{-j}, s_j) \tag{6.1}$$

If v satisfies this condition with $d = 1$, we say that v is an SOS valuation function.

Definition 28 (d-approximate strong submodular-over-signals valuations (d-strong-SOS valuations))**.** The valuation function $v(\mathbf{s})$ is a *d strong-SOS valuation* if for any j, $\delta \geq 0$,

$$\mathbf{s} = (s_1, \ldots, s_n) \quad \text{and} \quad \mathbf{s} = (s'_1, \ldots, s'_n)$$

[14] Weak diminishing returns submodularity was introduced in [Soma and Yoshida, 2015], where it's termed "diminishing returns submodularity".

such that s′ is smaller than or equal to s coordinate-wise, it holds that

$$d \cdot \left(v(\mathbf{s}'_{-j}, s'_j + \delta) - v(\mathbf{s}'_{-j}, s'_j)\right) \geq v(\mathbf{s}_{-j}, s_j + \delta) - v(\mathbf{s}_{-j}, s_j) \tag{6.2}$$

If v satisfies this condition with $d = 1$, we say that i's valuation functions are "strong-SOS".

Definition 29 (SOS-valuations settings)**.** We say that a mechanism design setting with interdependent valuations is an *SOS-valuations setting* or, equivalently, that the agents have SOS-valuations, in each of the following cases:

- Single parameter valuations (as in definition 15): for every i, the valuation function $v_i(\mathbf{s})$ is SOS.
- Combinatorial valuations with single-parameter signals (as in definition 23): for every i and T, the valuation function $v_{iT}(\mathbf{s})$ is SOS;
- Combinatorial valuations with multi-parameter signals (as in definition 24): for every i and T, $v_{iT}(\mathbf{s}_T)$ is SOS, where $\mathbf{s}_T = (s_{1T}, \ldots, s_{nT})$.

Similar definitions can be given for d-SOS valuation settings and d-strong-SOS valuation settings.

Finally, in section 6.5, we will specialize to the case of *separable* SOS valuations.

Definition 30 (Separable SOS valuations)**.** We say that a set of valuations as in Definition 24 are *separable SOS valuations* if for every agent i and subset T of items, $v_{iT}(\mathbf{s}_T)$ can be written as

$$v_{iT}(\mathbf{s}_T) = g_{-iT}(\mathbf{s}_{-iT}) + h_{iT}(s_{iT}),$$

where $g_{-iT}(\cdot)$ and $h_{iT}(\cdot)$ are both weakly increasing and $g_{-iT}(\mathbf{s}_{-iT})$ is itself an SOS valuation function.

Observation 25. *A separable SOS valuation function is itself an SOS valuation function.*

We can similarly define separable d-SOS valuations.

6.2.4 A useful fact about SOS valuations

Lemma 17. *Let $v : \mathbb{R}^{+n} \to \mathbb{R}^{+}$ be a d-SOS function. Let $A \subseteq [n]$ and $B = [n] \setminus A$. For any $\mathbf{s}_A, \mathbf{y}_A \in \mathbb{R}^{+|A|}$, and $\mathbf{s}_B, \mathbf{s}'_B \in \mathbb{R}^{+|B|}$ such that $\mathbf{s}_B$ is smaller than $\mathbf{s}'_B$ coordinate wise,*

$$d \cdot (v(\mathbf{s}_A + \mathbf{y}_A, \mathbf{s}_B) - v(\mathbf{s}_A, \mathbf{s}_B)) \geq v(\mathbf{s}_A + \mathbf{y}_A, \mathbf{s}'_B) - v(\mathbf{s}_A, \mathbf{s}'_B).$$

Proof. Let $i_1, i_2, \ldots, i_{|A|}$ be the elements of A. For $1 \leq j \leq |A|$, let $\mathbf{s}^j$ and $\mathbf{s}'^j$ denote the vectors

$$\begin{aligned} \mathbf{s}^j &= \Big((s_{i_1} + y_{i_1}), \ldots, (s_{i_j} + y_{i_j}), s_{i_{j+1}}, \ldots, s_{i_{|A|}}, \mathbf{s}_B\Big), \\ \mathbf{s}'^j &= \Big((s_{i_1} + y_{i_1}), \ldots, (s_{i_j} + y_{i_j}), s_{i_{j+1}}, \ldots, s_{i_{|A|}}, \mathbf{s}'_B\Big). \end{aligned}$$

Note that $\mathbf{s}^{|A|} = (\mathbf{s}_A + \mathbf{y}_A, \mathbf{s}_B)$, and $\mathbf{s}'^{|A|} = (\mathbf{s}_A + \mathbf{y}_A, \mathbf{s}'_B)$.

It follows from the d-SOS definition that for every $1 \leq j \leq |A|$,

$$d \cdot \left(v(\mathbf{s}^j) - v(\mathbf{s}^{j-1})\right) \geq v(\mathbf{s}'^j) - v(\mathbf{s}'^{j-1}), \tag{6.3}$$

where $\mathbf{s}^0 = (\mathbf{s}_A, \mathbf{s}_B)$ and $\mathbf{s}'^0 = (\mathbf{s}_A, \mathbf{s}'_B)$.

Summing Equation (6.3) for $j = 1, 2, \ldots, |A|$ proves the claim. □

6.3 The Key Lemma

The following is a key lemma which is used for both single parameter and combinatorial settings.

Lemma 18. *Let $v_i : \mathbb{R}^{+n} \to \mathbb{R}^{+}$ be a d-SOS function. Let A be a uniformly random subset of $[n] \setminus \{i\}$, and let $B := ([n] \setminus \{i\}) \setminus A$. It now holds that*

$$\mathbb{E}_A \left[v_i(\mathbf{s}_A, \mathbf{0}_B, s_i)\right] \geq \frac{1}{d+1} v_i(\mathbf{s}),$$

where the expectation is over the random choice of A.

Proof. We consider two equiprobable events,

- $A = S \subset [n] \setminus \{i\}$ is chosen as the random subset.

- $A = T = ([n] \setminus \{i\}) \setminus S$ is chosen as the random subset.

Normalize the valuations so that $v_i(\mathbf{s}) = 1$ and define $\alpha, \beta \in [0,1]$ such that

$$v_i(\mathbf{s}_S, \mathbf{0}_T, s_i) \;=\; \alpha, \qquad v_i(\mathbf{0}_S, \mathbf{s}_T, s_i) \;=\; \beta.$$

It follows that

$$\begin{aligned}\beta = v_i(\mathbf{0}_S, \mathbf{s}_T, s_i) \;&\geq\; v_i(\mathbf{0}_S, \mathbf{s}_T, s_i) - v_i(\mathbf{0}_S, \mathbf{0}_T, s_i) \\ &\geq\; (v_i(\mathbf{s}_S, \mathbf{s}_T, s_i) - v_i(\mathbf{s}_S, \mathbf{0}_T, s_i))/d \\ &=\; (1-\alpha)/d,\end{aligned}$$

where the first inequality follows from non-negativity of $v_i(\mathbf{0}_S, \mathbf{0}_T, s_i)$, and the second inequality follows from v_i being d-SOS and Lemma 17.

Similarly, we have that

$$\alpha \;\geq\; (1-\beta)/d \quad\Rightarrow\quad \beta \geq 1 - \alpha d;$$

It follows that

$$\alpha + \beta \geq \max\left(\alpha + \frac{1-\alpha}{d}, \alpha + 1 - \alpha d\right).$$

Solving for equality of the two terms, we get that $\alpha = 1/(d+1)$ which implies that

$$\alpha + \beta \geq \frac{2}{d+1}.$$

Partition the event space into pairs (S, T) that partition $[n] \setminus \{i\}$. For every such (S, T) pair, it follows that $v_i(\mathbf{s}_S, \mathbf{0}_T, s_i) + v_i(\mathbf{0}_S, \mathbf{s}_T, s_i) = \alpha + \beta \geq \frac{2}{d+1}$.

We conclude with the following, where the third line follows from the fact that there are $2^{n-1}/2$ such (S, T) pairs that partition $[n] \setminus \{i\}$:

$$\begin{aligned} \mathbb{E}_A\left[v_i(\mathbf{s}^A, \mathbf{0}_B, s_i)\right] &= \sum_{A \subseteq [n] \setminus \{i\}} \Pr[A] \cdot v_i(\mathbf{s}^A, \mathbf{0}_B, s_i) \\ &= \frac{1}{2^{n-1}} \cdot \sum_{A \subseteq [n] \setminus \{i\}} v_i(\mathbf{s}^A, \mathbf{0}_B, s_i) \\ &\geq \frac{1}{2^{n-1}} \cdot \frac{2^{n-1}}{2} \cdot \frac{2}{d+1} = \frac{1}{d+1}, \end{aligned}$$

as desired. □

6.4 Single-Parameter Valuations

In this section we describe the `Random Sampling Vickrey (RS-V) mechanism` that achieves a 4-approximation for single-parameter downward-closed environments with SOS valuations and a $2(d+1)$-approximation for d-SOS valuations. We then give a lower bound of 2 and $\sqrt{d}$ for SOS and d-SOS valuations respectively, even in the case of selling a single item.

Let $\mathcal{I} \subseteq 2^{[n]}$ be a downward-closed set system. We present a mechanism that serves only sets in $\mathcal{I}$ and gets a $2(d+1)$-approximation to the optimal welfare.

`Random Sampling Vickrey (RS-V):`

- Elicit bids $\tilde{\mathbf{s}}$ from the agents.
- Partition the agents into two sets, A and B, uniformly at random.
- For $i \in B$, let $w_i = v_i(\tilde{\mathbf{s}}_A, \tilde{s}_i, \mathbf{0}_{B\setminus\{i\}})$.
- Allocate to a set of bidders in

$$\text{argmax}_{S \in \mathcal{I}\,:\,S \subseteq B} \left\{ \sum_{i \in S} w_i. \right\}$$

Theorem 26. *For agents with SOS valuations, and for every downward-closed feasibility constraint $\mathcal{I}$, `RS-V` is an ex-post IC-IR mechanism that gives 4-approximation to the optimal welfare. For d-SOS valuations, the mechanism gives a $2(d+1)$-approximation to the optimal welfare.*

Proof. We first show the allocation is monotone in one's signal, and hence, by Proposition 24, the mechanism is ex-post IC-IR. Fix a random partition (A, B).

- Agents in A are never allocated anything and thus their allocation is weakly monotone in their signal.
- For an agent $i \in B$, increasing $\tilde{s}_i$ can only increase w_i, whereas it leaves w_j unchanged for all $j \in B \setminus \{i\}$. Thus, this only increases the weight of feasible sets (subsets of B in $\mathcal{I}$) that i belongs to. Therefore, increasing s_i can only cause i to go from being unallocated to being allocated.

For approximation, consider a set $S^* \in \text{argmax}_{S\in\mathcal{I}} \sum_{i\in S} v_i(\mathbf{s})$ that maximizes social welfare. For every $i \in S^*$, from the Key Lemma 18, we have that

$$\mathbb{E}_B[w_i \cdot \mathbf{1}_{i\in B}] = \mathbb{E}_B[v_i(\mathbf{s}_i, \mathbf{s}_A, \mathbf{0}_{B_{-i}}) \mid i \in B] \cdot Pr(i \in B) \geq \frac{v_i(\mathbf{s})}{d+1} \cdot \frac{1}{2}. \tag{6.4}$$

For every set B, the fact that $\mathcal{I}$ is downward-closed implies that $S^* \cap B \in \mathcal{I}$. Therefore, $S^* \cap B$ is eligible to be selected by `RS-V` as the allocated set of bidders. We have that the values of the bidders we allocate to are at least

$$\begin{aligned}\mathbb{E}_B\left[\max_{S\in\mathcal{I}:S\subseteq B}\sum_{i\in S} w_i\right] &\geq \mathbb{E}_B\left[\sum_{i\in S^*\cap B} w_i\right] = \mathbb{E}_B\left[\sum_{i\in S^*} w_i \cdot \mathbf{1}_{i\in B}\right] \\ &= \sum_{i\in S^*} \mathbb{E}_B\left[w_i \cdot \mathbf{1}_{i\in B}\right] \geq \sum_{i\in S^*} \frac{v_i(\mathbf{s})}{2(d+1)},\end{aligned}$$

as desired. Since the allocated bidders' true values at s are only higher than the proxy values w_i, this continues to hold.

□

We note that for the case of downward-closed feasibility constraints, even if the valuations satisfy single-crossing, there can be an $n-1$ gap between the optimal welfare and the welfare that the best deterministic mechanism can get. This is stated in Theorem 36 in Section C.2.

The following lower bounds, Theorem 27 show that even for a single item setting, one cannot hope to get a better approximation than 2 and $\Omega(\sqrt{d})$ for SOS and

d-SOS valuations respectively.The lower bounds apply to arbitrary randomized mechanisms[15].

Theorem 27. *No ex-post IC-IR mechanism (not necessarily universal) for selling a single item can get a better approximation than*

(a) a factor of 2 for SOS valuations.

(b) a factor of $\Omega(\sqrt{d})$ for d-SOS valuations.

Proof. Let $x_i(\mathbf{s})$ be the probability agent i is allocated at signal profile s. Notice that for every s, $\sum_i x_i(\mathbf{s}) \leq 1$, otherwise the allocation rule is not feasible.

(a) Consider the case where there are two agents, 1 and 2, $s_1 \in \{0, 1\}$ and agent 2 has no signal. The valuations are $v_1(0) = 1$, $v_1(1) = 1 + \epsilon$, $v_2(0) = 0$ and $v_2(1) = H$ for $H \gg 1 \gg \epsilon$. It is easy to see the valuations are SOS.

In order to get better than a 2-approximation at $s_1 = 0$, we must have $x_1(0) > 1/2$. By monotonicity, this forces $x_1(1) > 1/2$ as well, and hence $x_2(1) < 1/2$ by feasibility. This implies that the expected welfare when $s_1 = 1$ is $x_1(1)v_1(1) + x_2(1)v_2(1) < H/2 + 1$, while the optimal welfare when $s_1 = 1$ is H. For a large H, this approaches a 2-approximation. Note that this lower bound applies even given a known prior distribution on the signals in the event that we have a prior on the signals that satisfies: $\Pr[s_1 = 0] \cdot 1 = \Pr[s_1 = 1] \cdot H$.

(b) Consider the case where there are $n = \sqrt{d}$ agents and $s_i \in \{0, 1\}$ for every agent i. The valuation of agent i is

$$v_i(\mathbf{s}) = \begin{cases} \sum_{j \neq i} s_j + \epsilon \cdot s_i & \exists j \neq i \ : \ s_j = 0 \\ d + \epsilon \cdot s_i & s_j = 1 \, \forall j \neq i, \end{cases}$$

where $\epsilon \to 0$.

To see that the valuations are d-SOS, notice that whenever a signal s_j changes from 0 to 1, the valuation of agent $i \neq j$ increases by 1 *unless* all other signals

[15]A randomized mechanism takes as input the set of signals s and produces as output $x_i(\mathbf{s})$ and $p_i(\mathbf{s})$ for each agent i, where $x_i(\mathbf{s})$ is the probability that agent i wins and $p_i(\mathbf{s})$ is agent i's expected payment. Such a mechanism is ex-post IC (but not necessarily universally so) if and only if $x_i(s_i, \mathbf{s}_{-i})$ is monotonically increasing in s_i.

beside i's are already set to 1, in which case the valuation increases by $d-\sqrt{d}+2<d$. Consider valuation profiles $\mathbf{s^i}=(0_i,\mathbf{1}_{-i})$. Note that by monotonicity, for every truthful mechanism, it must be the case that $x_i(\mathbf{s^i})\leq x_i(\mathbf{1})$. Since any feasible allocation rule must satisfy $\sum_{i=1}^{\sqrt{d}} x_i(\mathbf{1})\leq 1$, then it must be the case there exists some agent i such that $x_i(\mathbf{1})\leq\frac{1}{\sqrt{d}}$, which by monotonicity implies that $x_i(\mathbf{s^i})\leq\frac{1}{\sqrt{d}}$. However, at profile $\mathbf{s^i}$, $v_i(\mathbf{s^i})=d$ while $v_j(\mathbf{s^i})=\sqrt{d}-2<\sqrt{d}$ for all $j\neq i$, so we get that the expected welfare of the mechanism at $\mathbf{s^i}$ is at most $x_i(\mathbf{s^i})\cdot d+(1-x_i(\mathbf{s^i}))\cdot\sqrt{d}\leq 2\sqrt{d}$, while the optimal welfare is d. Again, the lower bound also applies to the setting with known priors on the signals using a prior that satisfies: $\Pr[\mathbf{s^i}]=\Pr[\mathbf{s^j}]=\frac{1}{\sqrt{d}}$ for all i and j.

□

6.5 Combinatorial Auctions with Separable Valuations

In this section we present an ex-post IC-IR mechanism that gives $1/4$ of the optimal social welfare in any combinatorial auction setting with separable SOS valuations (as in Definition 30). The mechanism, that we call the `Random-sampling VCG` auction is a natural extension of the `Random-Sampling Vickrey (RS-V)` auction presented in Section 6.4. Note that unlike `RS-V`, here we need to explicitly define payments so that the obtained mechanism is ex-post IC-IR. We derive VCG-inspired payments which align the objective of the mechanism with that of the agents. Separability is used here, as without it, the payment term would have been affected by the agent's report (while with separability, only the allocation is affected by it).

`Random-Sampling VCG (RS-VCG):`

- Agents report their signals $\tilde{\mathbf{s}}$.
- Partition the agents into two sets A and B uniformly at random.
- For each agent $j\in B$ and bundle $T\subseteq[m]$, let

$$w_{jT}:=v_{jT}(\tilde{s}_{jT},\tilde{\mathbf{s}}_{AT},\mathbf{0}_{B_{-j}T})=g_{-jT}(\tilde{\mathbf{s}}_{AT},\mathbf{0}_{B_{-j}T})+h_{jT}(\tilde{s}_{jT}).$$

- Let the allocation be

$$\{T_i\}_{i\in B} \in \text{argmax}_{\{S_i\}_{i\in B}} \sum_{i\in B} w_{iS_i};$$

  i.e., $\{T_i\}_{i\in B}$ is the allocation that maximizes the "welfare" using w_{iT}'s.

- Set the payment for a winning agent $i \in B$ receiving set of goods T_i to be:

$$p_i(\tilde{\mathbf{s}}) := g_{-iT_i}(\tilde{\mathbf{s}}_{-iT_i}) - g_{-iT_i}(\tilde{\mathbf{s}}_{AT_i}, \mathbf{0}_{B_{-i}T_i}) - \sum_{j\in B\setminus\{i\}} w_{jT_j} + w_{-i},$$

  where

$$w_{-i} = \max_{\text{partitions } \{T'_j\}} \sum_{j\in B\setminus\{i\}} w_{jT'_j},$$

  that is, w_{-i} is the weight of the best allocation without agent i.

Since the w_{jT}'s do not depend on agent i's report (since i is in B), w_{-i} doesn't depend on agent i's report. Therefore, we can (and will) ignore this term when considering incentive compatibility below.

Note also that since the maximal partition guarantees that $w_{-i} \geq \sum_{j\in B\setminus\{i\}} w_{jT_j}$, and monotonicity of valuations in signals guarantees that $g_{-iT_i}(\tilde{\mathbf{s}}_{-i}) \geq g_{-iT_i}(\tilde{\mathbf{s}}_A, \mathbf{0}_{B_{-i}})$. Therefore, the payments $p_i(\tilde{\mathbf{s}})$ are always nonnegative.

Theorem 28. *`Random-Sampling VCG` is an ex-post IC-IR mechanism that gives a 4-approximation to the optimal social welfare for any combinatorial auction setting with separable SOS valuations.*

Proof. First we show that if the agents bid truthfully, then the mechanism gives a 4-approximation to social welfare. For every agent i and bundle T,

$$\mathbb{E}_B[w_{iT} \cdot \mathbf{1}_{i\in B}] = \mathbb{E}_B[v_{iT}(\mathbf{s}_{iT}, \mathbf{s}_{AT}, \mathbf{0}_{B_{-i}T}) \mid i \in B] \cdot Pr(i \in B) \geq \frac{v_{iT}(\mathbf{s}_T)}{2} \cdot \frac{1}{2}, \quad (6.5)$$

where the inequality follows by applying Lemma 18 with $d = 1$.

Let $S_1^*, \ldots, S_n^*$ be the true welfare maximizing allocation. Then,

$$\mathbb{E}_B \left[\max_{\text{partitions } \{T_i\}} \sum_{i \in B} w_{iT_i} \right] \geq \mathbb{E}_B \left[\sum_i w_{iS_i^*} \cdot \mathbf{1}_{i \in B} \right]$$
$$= \sum_i \mathbb{E}_B[w_{iS_i^*} \cdot \mathbf{1}_{i \in B}] \geq \frac{1}{4} \sum_i v_{iS_i^*}(\mathbf{s}_{S_i^*}),$$

where the last inequality follows by substituting S_i^* in T in Equation (6.5) for every i. Since $v_{iT}(\mathbf{s})$ is always at least w_{iT}, this proves the approximation ratio.

Next, we show that `RS-VCG` is universally ex-post IC. Fix a random partition (A, B). Suppose that when all agents bid truthfully

$$\{T_j^*\}_{j \in B} = \text{argmax}_{\text{partitions } \{T_j\}} \sum_{j \in B} w_{jT_j}.$$

Suppose that all agents but $i \in B$ bid truthfully and i bids s_i' instead of his true signal vector s_i. Let $\{T_j'\}_{j \in B}$ be the resulting allocation. Therefore, agent i's utility when reporting s_i' (after disregarding the w_{-i} term as mentioned above) is:

$$\begin{aligned}
v_{iT_i'}(\mathbf{s}) - p_i(\mathbf{s}_i', \mathbf{s}_{-i}) &= g_{-iT_i'}(\mathbf{s}_{-iT_i'}) + h_{iT_i'}(\mathbf{s}_{iT_i'}) - p_i(\mathbf{s}_i', \mathbf{s}_{-i}) \\
&= g_{-iT_i'}(\mathbf{s}_{-iT_i'}) + h_{iT_i'}(\mathbf{s}_{iT_i'}) - \left(g_{-iT_i'}(\mathbf{s}_{-iT_i'}) - g_{-iT_i'}(\mathbf{s}_{AT_i'}, \mathbf{0}_{B_{-i}T_i'}) - \sum_{j \in B \setminus \{i\}} w_{jT_j'} \right) \\
&= h_{iT_i'}(\mathbf{s}_{iT_i'}) + g_{-iT_i'}(\mathbf{s}_{AT_i'}, \mathbf{0}_{B_{-i}T_i'}) + \sum_{j \in B \setminus \{i\}} w_{jT_j'} \\
&= w_{iT_i'} + \sum_{j \in B \setminus \{i\}} w_{jT_j'} \; = \sum_{j \in B} w_{jT_j'} \\
&\leq \sum_{j \in B} w_{jT_j^*},
\end{aligned}$$

where $\sum_{j \in B} w_{jT_j^*}$ is i's utility for bidding truthfully.

Finally, we show that the mechanism is ex-post IR. Indeed, from above, agent i's utility when reporting truthfully (and without disregarding the w_{-i} term) is

$$v_{iT_i^*}(\mathbf{s}_{T_i^*}) - p_i(\mathbf{s}) = \sum_{j \in B} w_{jT_j^*} - w_{-i} = \sum_{j \in B} w_{jT_j^*} - \max_{\text{partitions } \{T_j'\}} \sum_{j \in B \setminus \{i\}} w_{jT_j'} \geq 0.$$

□

In the case of separable d-SOS valuations, the `Random-Sampling VCG` is an ex-post IC-IR mechanism that gives $2(d+1)$-approximation to the social welfare. The proof is identical to Theorem 28, except that Equation (6.5) is changed to

$$\mathbb{E}_B[w_{iT} \cdot \mathbf{1}_{i \in B}] \geq \frac{v_{iT}(\mathbf{s}_T)}{2(d+1)},$$

since we apply Lemma 18 with an arbitrary d.

Remark 29. *Theorem 28 is clearly analogous to the VCG mechanism for combinatorial auctions with private values. As with VCG for private values, in many cases, there is unlikely to be a polynomial time algorithm to compute allocations and payments. Exceptions include settings we know and love such as unit-demand auctions, additive valuations, etc.*

6.6 Combinatorial Auctions with Single-Dimensional Signals

In this section we consider combinatorial valuations (general combinatorial auctions) with single-dimensional signals (as given by Definition 23).

When the signal space of each agent is of size at most k, we present a mechanism that gets $(k+3)$-approximation for SOS valuations (see Section 6.6.1), and a mechanism that gets $(2\log_2 k + 4)$-approximation for strong-SOS valuations (Definition 28, see Section 6.6.2 for details regarding the mechanism). For d-SOS and d-strong-SOS valuations, the mechanism generalizes to give $O(dk)$- and $O(d^2 \log k)$-approximations respectively, as shown in Section C.3.

We first decompose the optimal welfare into two parts, OTHER and SELF. Each part will be covered by a corresponding mechanism. Let $T^* = \{T_i^*\}_{i \in [n]}$ be a welfare-maximizing allocation at signal profile s, and let $W^*(\mathbf{s})$ be the social welfare of T^* at

s. Consider the following decomposition:

$$\begin{aligned} W^*(\mathbf{s}) &= \sum_i v_{iT_i^*}(\mathbf{s}) \\ &= \sum_i v_{iT_i^*}(\mathbf{s}_{-i}, 0_i) + \sum_{i\,:\,s_i>0} \big(v_{iT_i^*}(\mathbf{s}) - v_{iT_i^*}(\mathbf{s}_{-i}, 0_i)\big) \\ &\le \sum_i v_{iT_i^*}(\mathbf{s}_{-i}, 0_i) + \sum_{i\,:\,s_i>0} \big(v_{iT_i^*}(\mathbf{0}_{-i}, s_i) - v_{iT_i^*}(\mathbf{0})\big) \qquad (6.6) \\ &\le \underbrace{\sum_i v_{iT_i^*}(\mathbf{s}_{-i}, 0_i)}_{\text{OTHER}} + \underbrace{\sum_{\ell=1}^{k-1} \sum_{i\,:\,s_i=\ell} v_{iT_i^*}(\mathbf{0}_{-i}, s_i)}_{\text{SELF}}, \qquad (6.7) \end{aligned}$$

where Equation (6.6) follows from the definition of submodularity (and therefore, also follows the definition of strong-submodularity). The last inequality follows from the non-negativity of $v_{iT_i^*}(\mathbf{0})$. The first term in the decomposition represents the contribution of others' signals to one's value from his allocated bundle, while the second term represents one's contribution to his own value. Each of these terms will be targeted using a different mechanism. Whereas the OTHER term will be targeted using the same mechanism in both the SOS and strong-SOS cases, the SELF term will be treated differently.

6.6.1 $(k+3)$-approximation for SOS valuations

Suppose $s_i \in \{0, 1, \ldots, k-1\}$ for all i. The mechanism is as follows:

Mechanism `k signals High-Low (k-HL):`
With probability $p_{RT} = \frac{k-1}{k+3}$, run `Random Threshold`; otherwise, run `Random Sampling`, as described below:
Mechanism `Random Threshold`

- Choose a random threshold ℓ uniformly in $\{1, \ldots, k-1\}$.
- Let $N_{\geq \ell} = \{i \ : \ s_i \geq \ell\}$ be the "high" agents; i.e., agents with signal at least ℓ, and let $N_{<\ell} = [n] \setminus N_{\geq \ell}$ be the "low" agents.
- For every high agent $i \in N_{\geq \ell}$ and bundle T, let $\bar{v}_{iT} := v_{iT}(\mathbf{s}_{N_{<\ell}}, \ell_{N_{\geq \ell}})$
- For every low agent $i \in N_{<\ell}$ and bundle T, let $\bar{v}_{iT} := 0$.

- Let the allocation be

$$\bar{T} \in \mathrm{argmax}_{S=\{S_i\}_{i\in N_{\geq \ell}}} \sum_{i\in N_{\geq \ell}} \bar{v}_{iS_i}.$$

  (*i.e.*, the allocation that maximizes the "welfare" of high agents using values $\bar{v}_{iT}$.)

- Agent i that receives bundle $\bar{T}_i$ pays $v_{i\bar{T}_i}(\mathbf{s}_{-i}, s_i = \ell - 1)$.

Mechanism `Random Sampling`

- Split the agents into sets A and B uniformly at random.
- For each $i \in B$ and bundle T, let $\tilde{v}_{iT} := v_{ij}(\mathbf{s}_A, \mathbf{0}_B)$.
- For each $i \in A$ and bundle T, let $\tilde{v}_{iT} := 0$.
- Let the allocation be

$$\tilde{T} \in \mathrm{argmax}_{S=\{S_i\}_{i\in B}} \sum_{i\in B} \tilde{v}_{iS_i}.$$

  (*i.e.*, the allocation that maximizes the "welfare" of agents in B using values $\tilde{v}_{iT}$.)

- Charge no payments.

The `k-HL` mechanism is a random combination of two mechanisms: `Random Threshold` approximates the welfare contribution of the bidders' signals to their own value (the SELF term); `Random Sampling` approximates the welfare contributions of the bidders' signals to other bidders' values (the OTHER term). We wish to prove the following theorem.

Theorem 30. *For every combinatorial auction setting with SOS valuations, single-dimensional signals, and signal space of size k, i.e. $s_i \in \{0, 1, \ldots, k-1\}\, \forall i$, mechanism k-`HL` is an ex-post IC-IR mechanism that gives $(k+3)$-approximation to the optimal social welfare.*

We first argue that the mechanism is ex-post IC-IR. Proof of ex-post IC-IR. `Random`

`Sampling` is ex-post IC-IR since the agents that might receive items (agents in B) cannot change the allocation since their signals are ignored (and they pay nothing).

As for `Random Threshold`, consider a threshold ℓ chosen by the mechanism. If the agent's signal is below ℓ and the agent reports ℓ or above, then his payment, if allocated bundle T is $v_{iT}(\mathbf{s}_{-i}, s_i = \ell - 1) \geq v_{iT}(\mathbf{s})$; *i.e.*, the agent's utility is non-positive. Bidding a different value below ℓ will grant the agent no items. If his value is ℓ or above, then bidding a different signal above ℓ will result in the same outcome, since the sets $N_{\geq \ell}$ and $N_{<\ell}$ remain the same. If he bids a signal below ℓ, then he won't receive any item, and his utility will be 0, while bidding his true signal will result in non-negative utility. $\square$ In Lemma 20, we prove that `Random Sampling` covers the OTHER component of the social welfare, and in Lemma 19, we show that `Random Threshold` covers the SELF component.

Lemma 19. *For SOS valuations, the* `Random Threshold` *mechanism gives a* $(k-1)$*-approximation to the SELF component of the optimal social welfare.*

Proof. Consider a threshold $\ell \in \{1, \ldots, k-1\}$ chosen in `Random Threshold`. Whenever ℓ is chosen, we have that

$$\sum_{i\,:\,s_i=\ell} \bar{v}_{iT_i^*} \;=\; \sum_{i\,:\,s_i=\ell} v_{iT_i^*}(\mathbf{s}_{N_{<\ell}}, \ell_{N_{\geq\ell}}) \;\geq\; \sum_{i\,:\,s_i=\ell} v_{iT_i^*}(\mathbf{0}_{-i}, s_i).$$

Since `Random Threshold` chooses an allocation $\bar{T} = \{\bar{T}_i\}_{i \in N_{\geq \ell}}$ that maximizes the welfare under $\bar{v}_{iT}$'s, the value of the allocation is only larger than the left expression above. Because $v_{i\bar{T}_i}(\mathbf{s}) \geq \bar{v}_{i\bar{T}_i}$, we get that if ℓ was chosen, which happens with probability $\frac{1}{k-1}$, the welfare achieved is at least $\sum_{i\,:\,s_i=\ell} v_{iT_i^*}(\mathbf{0}_{-i}, s_i)$. Therefore, the welfare from running `Random Threshold` is at least

$$\sum_{\ell=1}^{k-1} \frac{1}{k-1} \sum_{i\,:\,s_i=\ell} v_{iT_i^*}(\mathbf{0}_{-i}, s_i), \geq \frac{\mathsf{SELF}}{k-1}.$$

$\square$

Lemma 20. *For SOS valuations, the* `Random Sampling` *mechanism gives a* 4*-approximation to the OTHER component of the optimal social welfare.*

Proof. Consider a set T. Using an application of the Key Lemma 18 with respect to $v_{iT}(\mathbf{s}_{-i}, 0_i)$, we see that

$$\mathbb{E}_{A,B}[\tilde{v}_{iT}] \geq \Pr[i \in B] \cdot \mathbb{E}_{A,B}[\tilde{v}_{iT} \mid i \in B] = \frac{1}{2}\mathbb{E}_{A,B\setminus i}[\tilde{v}_{iT} \mid i \in B] \geq \frac{1}{4} v_{iT}(\mathbf{s}_{-i}, 0_i). \quad (6.8)$$

Therefore, the expected weight of the allocation $\{T_i^*\}_{i\in[n]}$ using weights $\tilde{v}_{iT}$'s is

$$\mathbb{E}_{A,B}\left[\sum_i \tilde{v}_{iT_i^*}\right] = \sum_i \mathbb{E}_{A,B}\left[\tilde{v}_{iT_i^*}\right] \geq \sum_i \frac{1}{4} v_{iT_i^*}(\mathbf{s}_{-i}, 0_i) = \frac{\mathsf{OTHER}}{4}.$$

Since the mechanism chooses the optimal allocation according to the $\tilde{v}_{iT}$'s, its weight can only be larger. Moreover, since $\tilde{v}_{iT} = v_{iT}(\mathbf{s}_{-i}, 0) \leq v_{iT}(\mathbf{s})$, the welfare achieved by the mechanism is at least $\frac{\mathsf{OTHER}}{4}$, as desired. □

We conclude by proving the claimed approximation ratio.

Proof of approximation. According to Lemma 19, `Random Threshold` approximates SELF to a factor of $k-1$. According to Lemma 20 that `Random Sampling` approximates OTHER to a factor of 4. Therefore, running `Random Threshold` with probability p_{RT} and `Random Sampling` with probability $1 - p_{RT}$ yields a welfare of

$$\begin{aligned} p_{RT}\frac{\mathsf{SELF}}{k-1} + (1-p_{RT})\frac{\mathsf{OTHER}}{4} &= \frac{k-1}{k+3} \cdot \frac{\mathsf{SELF}}{k-1} + \frac{4}{k+3} \cdot \frac{\mathsf{OTHER}}{4} \\ &= \frac{\mathsf{SELF} + \mathsf{OTHER}}{k+3} \geq \frac{W^*(\mathbf{s})}{k+3}, \end{aligned}$$

where the inequality follows Equation (6.7). □

6.6.2 $O(\log k)$-Approximation with Strong-SOS Valuations

Strong-SOS valuations means the effect on the valuation is concave in one's own signal. This allows us to use a bucketing technique in order to give an $O(\log k)$-approximation to the SELF component in the decomposition depicted by Equation (6.7).

Consider the SELF term in Equation (6.7). We can bound this term as follows:

$$\begin{aligned}
\mathsf{SELF} &= \sum_{\ell=1}^{k-1} \sum_{i\,:\,s_i=\ell} v_{iT_i^*}(\mathbf{0}_{-i}, s_i) \\
&= \sum_{\ell=1}^{\log_2 k} \sum_{i\,:\,2^{\ell-1}\le s_i<2^{\ell}} v_{iT_i^*}(\mathbf{0}_{-i}, s_i) \\
&\le \sum_{\ell=1}^{\log_2 k} \sum_{i\,:\,2^{\ell-1}\le s_i<2^{\ell}} v_{iT_i^*}(\mathbf{0}_{-i}, {2^{\ell-1}}_i),
\end{aligned} \tag{6.9}$$

where the inequality follows the definition of strong-SOS valuations.

We introduce mechanism `Random Bucket` to give an $O(\log k)$-approximation to the upper bound in Equation (6.9).

Mechanism `Random Bucket`:

- choose ℓ uniformly in $\{1, \ldots, \log_2 k\}$.
- Let $N_{B_\ell} = \{i \ :\ \text{such that } s_i \geq 2^{\ell-1}\}$ be the agents with signal at least $2^{\ell-1}$ and $N_{\neg B_\ell} = [n] \setminus N_{B_\ell}$.
- For $i \in N_{B_\ell}$ and bundle T, let $\bar{v}_{iT} := v_{iT}(\mathbf{s}_{N_{\neg B_\ell}}, \mathbf{2^{\ell-1}}_{N_{B_\ell}})$ (and $\bar{v}_{iT} := 0$ for $i \in N_{\neg B_\ell}$).
- Let the allocation be

  $$\bar{T} \in \mathrm{argmax}_{S=\{S_i\}_{i\in N_{B_\ell}}} \sum_{i\in N_{B_\ell}} \bar{v}_{iS_i}.$$

  (*i.e.*, the allocation that maximizes the "welfare" of high agents using values $\bar{v}_{iT}$.)
- Agent i that receives bundle $\bar{T}_i$ pays $v_{i\bar{T}_i}(\mathbf{s}_{-i}, s_i = 2^{\ell-1} - 1)$.

We show the following approximation guarantee regarding `Random Bucket`.

Lemma 21. *For strong-SOS valuations, the* `Random Bucket` *mechanism is ex-post IC-IR and gives a* $2\log_2 k$ *approximation to the* SELF *component of the optimal social welfare.*

Proof. The proof of ex-post IC-IR is identical to that of mechanism `Random Threshold`, as both are threshold-based mechanisms. The proof of the approximation guarantee is also very similar to that of `Random Threshold`.

Consider a threshold $2^{\ell-1}$ for $\ell \in \{1, \ldots, k-1\}$ chosen in `Random Bucket`. Whenever ℓ is chosen, we have that

$$\sum_{i\,:\,2^{\ell-1} \leq s_i < 2^{\ell}} \bar{v}_{iT_i^*} = \sum_{i\,:\,2^{\ell-1} \leq s_i < 2^{\ell}} v_{iT_i^*}(\mathbf{s}_{N_{\neg B_\ell}}, \mathbf{2^{\ell-1}}_{N_{B_\ell}}) \geq \sum_{i\,:\,2^{\ell-1} \leq s_i < 2^{\ell}} v_{iT_i^*}(\mathbf{0}_{-i}, 2^{\ell-1}{}_i).$$

Since `Random Bucket` chooses an allocation that maximizes the $\bar{v}_{iT}$'s, the value of the allocation is only larger. Because $v_{i\bar{T}_i}(\mathrm{s}) \geq \bar{v}_{i\bar{T}_i}$, we get that if ℓ was chosen, which happens with probability $\frac{1}{\log_2 k}$, the welfare achieved is at least $\sum_{i\,:\,2^{\ell-1} \leq s_i < 2^{\ell}} v_{iT_i^*}(\mathbf{0}_{-i}, s_i)$. Therefore, the welfare from running `Random Bucket` is at least

$$\sum_{\ell=1}^{\log_2 k} \frac{1}{\log_2 k} \sum_{i\,:\,2^{\ell-1} \leq s_i < 2^{\ell}} v_{iT_i^*}(\mathbf{0}_{-i}, s_i), \geq \frac{\mathsf{SELF}}{2\log_2 k}.$$

□

Mechanism k`-signals Strong-SOS` (k`-SS`) runs `Random Bucket` with probability $p_{RB} = \frac{log_2 k}{log_2 k+2}$ and mechanism `Random Sampling` with probability $1 - p_{RB}$.

Theorem 31. *For every combinatorial auction with single-dimensional signals with strong-SOS valuations and signal space of size k, i.e. $s_i \in \{0, 1, \ldots, k-1\}\, \forall i$, mechanism k`-SS` is ex-post IC-IR, and gives $(2\log_2 k + 4)$-approximation to the optimal social welfare.*

Proof. We already established that both `Random Bucket` and `Random Sampling` are ex-post IC-IR, hence k`-SS` is ex-post IC-IR as well. As for the approximation, according to Lemma 21, with probability p_{RB} we get $2\log_2 k$-approximation to SELF, and according to Lemma 20, with probability $1 - p_{RB}$ we get a 4-approximation to OTHER. Overall, the expected welfare is at least

$$\begin{aligned} p_{RB}\frac{\mathsf{SELF}}{2\log_2 k} + (1-p_{RB})\frac{\mathsf{OTHER}}{4} &= \frac{\mathsf{SELF} + \mathsf{OTHER}}{2\log_2 k + 4} \\ &\geq \frac{W^*}{2\log_2 k + 4}, \end{aligned}$$

as desired.

□

6.7 Open Problems

Our analysis and results suggest many open problems:

- For combinatorial auctions with multi-dimensional signals: is separability a necessary condition for achieving constant approximation to welfare? This problem is open even for single-dimensional signals, and even for “simple” combinatorial valuations, such as unit-demand.
- For single-parameter SOS valuations, downward closed feasibility, and single-dimensional signals, closing the gap between $1/4$ and $1/2$ is open.
- The exact same gap applies for combinatorial, separable-SOS valuations with multi-dimensional signals.
- How does the distinction between SOS and strong-SOS affect the problems above, if at all?
- When considering the relaxation of SOS valuations to d-SOS valuations, there is a gap between the positive and negative results with respect to the dependence on d.

More generally, what other classes of valuations give rise to approximately efficient mechanisms in settings with interdependent valuations?

A Missing Details from: The FedEx Problem

A.1 Proof of strong duality

Theorem 32. *Let $a_i(\cdot)$, $b_i(\cdot)$, $\lambda_i(\cdot)$, $\alpha_i(\cdot)$ be functions feasible for the primal and dual, satisfying all the conditions from Section sec:CS. Then they are optimal.*

Proof. First, we prove weak duality. For any feasible primal and dual:

$$\int_0^H \sum_{i=1}^m b_i(v)\, dv \tag{A.1}$$

$$= \int_0^H \sum_{i=1}^m \left(1 \cdot b_i(v) + 0 \cdot [\lambda_i(v) + \alpha_i(v)]\right)\, dv. \tag{A.2}$$

Applying primal feasibility, we see that this quantity is

$$\geq \int_0^H \sum_{i=1}^m \left(a_i(v) b_i(v)) - a_i'(v)\lambda_i(v) + \left[\int_0^v a_i(x) - a_{i+1}(x) dx\right] \alpha_i(v) \right)\, dv. \tag{A.3}$$

We rewrite this expression using the following.

- Applying integration by parts,using the facts that $\lambda_i(\cdot)$ is continuous (Condition (3.11)) and equal to 0 at any point that $a_i'(v) = \infty$,[1] we get

  $$-\int_0^H a_i'(v)\lambda_i(v)\, dv = -a_i(v)\lambda_i(v)\Big|_0^H + \int_0^H a_i(v)\lambda_i'(v)\, dv = \int_0^H a_i(v)\lambda_i'(v)\, dv,$$

  since $a_i(0) = 0$ and $\lambda_i(H) = 0$.

- Second, interchanging the order of integration, we get

  $$\int_0^H \int_0^v \left[a_i(x) - a_{i+1}(x) dx\right] \alpha_i(v)\, dv = \int_0^H \left(a_i(v) \int_v^H \alpha_i(x)\, dx - a_{i+1}(v) \int_v^H \alpha_i(x)\, dx \right)\, dv.$$

[1] $a_i'(v)$ can be ∞ at only countably many points.

Combining these shows that (A.3) equals

$$\int_0^H \left(\sum_{i=1}^m a_i(v) \left[b_i(v) + \lambda_i'(v) + \int_v^H \alpha_i(x) - \int_v^H \alpha_{i-1}(x)\, dx \right] \right) dv$$
$$\geq \int_0^H \sum_{i=1}^H a_i(v)\gamma_i(v)\, dv \tag{A.4}$$

where the last inequality is dual feasibility. (Note that $\alpha_0(\cdot) = \alpha_m(\cdot) = 0$.)

Comparing (A.1) and (A.4) yields weak duality, i.e., $\sum_i \int_0^H b_i(v)\, dv \geq \sum_i \int_0^H a_i(v)\gamma_i(v)\, dv$.

If the conditions (3.11)-(3.17) hold, we also have strong duality and hence optimality: To show that (A.2) = (A.3), observe that

- (4.2) $a_i(v) < 1$ implies that $b_i(v) = 0$;
- (4.3) $a_i'(v) > 0$ implies that $\lambda_i(v) = 0$.
- (4.4) $\int_0^v (a_{i+1}(x) - a_i(x))dx > 0$ implies that $\alpha_i(v) = 0$ for $i = 1, \ldots, n-1$.

Finally, (A.4) is an equality rather than an inequality because of conditions (4.5)-(3.17). □

B Missing Details from: Proportional Complementarities

B.1 Improved Additive Bound

We now give a proof which improves the 6-approximation by Babaioff et al. [2014] to 5.382.

Theorem 33. *For any $a > 0$,*

$$\text{OPT} \leq \left(2 + \frac{2}{a^2}\right) \cdot \text{BREV} + (a+1) \cdot \text{SREV}.$$

In particular, if we choose $a = \sqrt[3]{4}$, then

$$\text{OPT} \leq \left(3 + \frac{3}{2}\sqrt[3]{4}\right) \cdot \max\{\text{SREV}, \text{BREV}\} \leq 5.382 \cdot \max\{\text{SREV}, \text{BREV}\}.$$

Proof of Theorem 33. We improve the analysis used in Cai et al. [2016], where they obtain an upper bound on OPT using duality. They further partition the upper bound into three parts:

$$\text{OPT} \leq \text{SINGLE} + \text{TAIL} + \text{CORE}.$$

The first term SINGLE is upper bounded by SREV. The second term TAIL is also upper bounded by SREV, but the first thing we show is that it can also be upper bounded by BREV.

Let item j's value t_j be drawn from F_j independently, and $f_j(v_j)$ be the probability that $t_j = v_j$. Following the notation of Cai et al. [2016], we use R to denote SREV, and TAIL is defined as follows.

$$\text{TAIL} = \sum_{j\in[m]} \sum_{t_j > R} f_j(t_j) \cdot t_j \cdot \Pr_{t_{-j}\sim F_{-j}}[\exists \ell \neq j,\ t_\ell \geq t_j].$$

This quantity is the expected value above r from all but the highest item. Note that for any j and any t_j, selling the grand bundle at a price of t_j earns revenue at least $t_j \cdot \Pr_{t_{-j}\sim F_{-j}}[\exists \ell \neq j,\ t_\ell \geq t_j]$. Hence,

$$\textsc{tail} \leq \textsc{brev} \cdot \left(\sum_{j\in[m]} \sum_{t_j > R} f_j(t_j) \right) \leq \textsc{brev}.$$

The second inequality is because R_j is the optimal revenue for selling only item j, and $R_j \geq R \cdot \Pr[t_j \geq R]$, thus $\sum_{t_j > R} f_j(t_j) \leq R_j / R$; also, $R = \sum_j R_j$.

Next, we improve the analysis of the term $\textsc{core}$. In Cai et al. [2016], $\textsc{core}$ is upper bounded by $2\textsc{brev} + 2\textsc{srev}$. They make use of Chebyshev's inequality to obtain this bound. We improve their analysis using a tighter inequality due to Cantelli.

The $\textsc{core}$ is defined as follows.

$$\textsc{core} = \sum_{j\in[m]} \sum_{t_j \leq R} f_j(t_j) \cdot t_j = \mathbb{E}_{t\sim F}\left[\sum_{j\in[m]} t_j \cdot \mathbb{1}[t_j \leq R] \right]$$

It is shown in Cai et al. [2016] that $\text{Var}_{t\sim F}\left[\sum_{j\in[m]} t_j \cdot \mathbb{1}[t_j \leq R]\right] \leq 2R^2$. Now we state Cantelli's inequality:

Theorem 34 (Cantelli's Inequality)**.** *For any real valued random variable X and any positive number τ,*

$$\Pr\left[X \geq \mathbb{E}[X] - \tau\right] \geq \frac{\tau^2}{\tau^2 + \text{Var}[X]}.$$

We define the random variable $V = \sum_{j\in[m]} t_j \cdot \mathbb{1}[t_j \leq R]$ and apply Cantelli's inequality to V with $\tau = aR$.

$$\Pr[V \geq \textsc{core} - aR] \geq \frac{a^2R^2}{\text{Var}[V] + a^2R^2} \geq \frac{a^2}{2 + a^2}.$$

The last inequality is because $\text{Var}[V] \leq 2R^2$. Therefore, $\textsc{brev} \geq (\textsc{core} - aR) \cdot \frac{a^2}{2+a^2}$, which implies $\textsc{core} \leq (1 + \frac{2}{a^2}) \cdot \textsc{brev} + a \cdot \textsc{srev}$. Combining our new analysis for the $\textsc{tail}$ and the $\textsc{core}$, we obtain the new bound. □

C Missing Details from: Combinatorial Auctions with Interdependent Valuations

C.1 Unit-Demand Valuations with Single-Crossing

Whereas single-crossing is a strong enough condition to implement the fully efficient mechanism in a variety of single-parameter environment, generalizations of this condition fail even in the simplest multi-parameter environments. We consider the case where bidders are unit demand and each bidder has a scalar as a signal. We define single-crossing for this setting as follows.

Definition 31 (Single-crossing for unit-demand valuations)**.** A valuation profile $\mathbf{v}$ is said to be single crossing if for every agent i, signals $\mathbf{s}_{-i}$, item j and agent ℓ,

$$\frac{\partial}{\partial s_i} v_{ij}(\mathbf{s}_{-i}, s_i) \geq \frac{\partial}{\partial s_i} v_{\ell j}(\mathbf{s}_{-i}, s_i). \tag{C.1}$$

In this section, we show that in the case two non-identical items are for sale, and the valuations are unit demand and satisfy single-crossing as defined in Equation (C.1), any truthful mechanism is bounded away from achieving full efficiency.

In order to give the lower bound, we first give a characterization of ex-post IC and IR mechanisms in multi-dimensional environments in interdependent values settings (Section C.1.1). We then turn to prove the lower bound (Section C.1.2).

C.1.1 Cycle Monotonicity

In the IPV model, Rochet [1987] introduced cycle monotonicity as a necessary and sufficient condition on the allocation to be implementable in dominant strategies (DSIC) for multidimensional environments. It was noticed that a straightforward analogue holds for the IDV value model, for ex-post implementability (EPIC) (in Vohra [2007], this fact is stated without a proof).

Fix a feasible allocation rule $\mathbf{x} = \{x_i\}_{i\in[n]}$, where $x_{iT}(\mathbf{s})$ is the probability agent i receives a bundle T under bid profile $\mathbf{s}$. For each agent i, consider the graph $G_i^{\mathbf{x}}$

where there is a vertex for each signal profile $\mathbf{s}$, and there is a directed edge from $\mathbf{s}$ to $\mathbf{t}$ if $\mathbf{s}_{-i} = \mathbf{t}_{-i}$. The weight of edge $(\mathbf{s}, \mathbf{t})$ is

$$w(\mathbf{s}, \mathbf{t}) = \mathbb{E}_{T \sim x_i(\mathbf{s})}[v_{iT}(\mathbf{s})] - \mathbb{E}_{T \sim x_i(\mathbf{t})}[v_{iT}(\mathbf{s})] = \sum_{T \subseteq [m]} x_{iT}(\mathbf{s}) v_{iT}(\mathbf{s}) - \sum_{T \subseteq [m]} x_{iT}(\mathbf{t}) v_{iT}(\mathbf{s}).$$

The following theorem states that a necessary and sufficient condition for ex-post implementability of $\mathbf{x}$ is that for every agent i, every directed cycle in $G_i^{\mathbf{x}}$ is non-negative. The proof is a straightforward adjustment of the original proof in Rochet [1987], and is given below for completeness.

Theorem 35. *The allocation rule $\mathbf{x}$ is implementable by an ex-post IC mechanism if and only if for every agent i, all directed cycles in $G_i^{\mathbf{x}}$ have non-negative weight.*

Proof. We first show that if the allocation rule is implementable, then there are no negative cycles. Fix some payment rule $\mathbf{p} = \{p_i\}_{i \in [n]}$, where $p_i(\mathbf{s})$ is the payment of agent i under bid profile $\mathbf{s}$. Let $\mathbf{s}_{-i}$ be the real signals of all bidders except i, and consider a cycle $\mathbf{s}^1 \to \mathbf{s}^2 \to \ldots \to \mathbf{s}^\ell \to \mathbf{s}^1$ in $G_i^{\mathbf{x}}$, where $\mathbf{s}^t = (\mathbf{s}_{-i}, s_i = \zeta_t)$ for $t \in [\ell]$. Since $(\mathbf{x}, \mathbf{p})$ is an ex-post IC mechanism, for every true signal $s_i = s$, agent i is at least as well off bidding s than any other bid s'. We get that

$$\begin{aligned}
\mathbb{E}_{T \sim x_i(\mathbf{s}^1)}[v_{iT}(\mathbf{s}^1)] - p_i(\mathbf{s}^1) &\geq \mathbb{E}_{T \sim x_i(\mathbf{s}^2)}[v_{iT}(\mathbf{s}^1)] - p_i(\mathbf{s}^2) \\
&\vdots \\
\mathbb{E}_{T \sim x_i(\mathbf{s}^{\ell-1})}[v_{iT}(\mathbf{s}^{\ell-1})] - p_i(\mathbf{s}^{\ell-1}) &\geq \mathbb{E}_{T \sim x_i(\mathbf{s}^\ell)}[v_{iT}(\mathbf{s}^{\ell-1})] - p_i(\mathbf{s}^\ell) \\
\mathbb{E}_{T \sim x_i(\mathbf{s}^\ell)}[v_{iT}(\mathbf{s}^\ell)] - p_i(\mathbf{s}^\ell) &\geq \mathbb{E}_{T \sim x_i(\mathbf{s}^1)}[v_{iT}(\mathbf{s}^\ell)] - p_i(\mathbf{s}^1)
\end{aligned}$$

Summing over the above inequalities and using the convention that $\ell + 1 = 1$, we get that

$$\begin{aligned}
& \sum_{j=1}^{\ell} \mathbb{E}_{T \sim x_i(\mathbf{s}^j)}[v_{iT}(\mathbf{s}^j)] - \sum_{j=1}^{\ell} p_i(\mathbf{s}^j) \geq \sum_{j=1}^{\ell} \mathbb{E}_{T \sim x_i(\mathbf{s}^{j+1})}[v_{iT}(\mathbf{s}^j)] - \sum_{j=1}^{\ell} p_i(\mathbf{s}^j) \\
\iff \quad & \sum_{j=1}^{\ell} \left(\mathbb{E}_{T \sim x_i(\mathbf{s}^j)}[v_{iT}(\mathbf{s}^j)] - \mathbb{E}_{T \sim x_i(\mathbf{s}^{j+1})}[v_{iT}(\mathbf{s}^j)] \right) \geq 0,
\end{aligned}$$

where the LHS of the last inequality is exactly the weight of the cycle.

We now show how to compute payments that implement a given allocation rule x that induces no negative cycles for any i and $G_i^{\mathbf{x}}$. Given $G_i^{\mathbf{x}}$, one can compute payments as follows.

- Add a dummy node d with edges of weight 0 to all nodes in $G_i^{\mathbf{x}}$.
- For every node s of $G_i^{\mathbf{x}}$, let $\delta(\mathbf{s})$ be the distance of the shortest path from d to s.
- Set $p_i(\mathbf{s}) = -\delta(\mathbf{s})$.

Fix signals of the other players $\mathbf{s}_{-i}$. Let s be player i's true signal and s' be some other possible signal for i. Denote $\mathbf{s} = (\mathbf{s}_{-i}, s)$ and $\mathbf{s}' = (\mathbf{s}_{-i}, s')$. Consider the nodes s and s′ in $G_i^{\mathbf{x}}$. Since $\delta(\mathbf{s}')$ is the length of the shortest path from d, it must be that

$$\delta(\mathbf{s}') \leq \delta(\mathbf{s}) + w(\mathbf{s}, \mathbf{s}'),$$

where $w(\mathbf{s}, \mathbf{s}')$ is the weight of the edge from s to s′. Substituting $w(\mathbf{s}, \mathbf{s}') = \mathbb{E}_{T\sim x_i(\mathbf{s})}[v_{iT}(\mathbf{s})] - \mathbb{E}_{T\sim x_i(\mathbf{s}')}[v_{iT}(\mathbf{s})]$, $p_i(\mathbf{s}) = -\delta(\mathbf{s})$, and $p_i(\mathbf{s}') = -\delta(\mathbf{s}')$, we get

$$\mathbb{E}_{T\sim x_i(\mathbf{s})}[v_{iT}(\mathbf{s})] - p_i(\mathbf{s}) \geq \mathbb{E}_{T\sim x_i(\mathbf{s}')}[v_{iT}(\mathbf{s})] - p_i(\mathbf{s}'),$$

as desired. □

C.1.2 Lower Bounds for Deterministic and Randomized Mechanisms

Lemma 22. *There exists a setting with two items and two agents with unit-demand and single crossing valuations, such that no deterministic truthful mechanism achieves more than $1/2$ of the optimal welfare.*

Proof. Consider the setting depicted in Figure C.1, with two agents, 1 and 2, and two items, a and b. $s_1 \in \{0, 1\}$ and s_2 is fixed. The values at $s_1 = 0$ are

$$v_{1a}(0) = 1, v_{1b}(0) = 0, v_{2a}(0) = 0, v_{1b}(0) = 1,$$

and at $s_1 = 1$ are

$$v_{1a}(1) = 1 + H + \epsilon, v_{1b}(1) = H, v_{2a}(1) = H, v_{1b}(1) = 1,$$

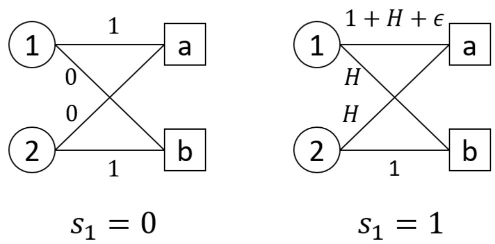


Figure C.1: An instance with unit-demand single-crossing valuations where no deterministic truthful allocation achieves more than a half of the optimal welfare.

for some arbitrarily large H and a sufficiently small ϵ. One can easily verify that the valuations satisfy Equation (C.1), and hence single crossing; indeed, when agent 1's signal increases, the valuation of agent 1 for each one of the item increases by more than the change in agent 2's valuation.

We show that no deterministic truthful mechanism can get better than 2-approximation. In order to get better than 2-approximation, the mechanism must allocate item a to agent 1 and item b to bidder 2 at signal $s_1 = 0$. At $s_1 = 1$, allocating item b to agent 1 and item a to agent 2 obtains a welfare of $2H$, while any other allocation obtains at most a welfare of $H + 2 + \epsilon$. Since H can be arbitrarily large, one must allocate item b to agent 1 and item a to agent 2 at signal $s_1 = 1$ in order to get an approximation ratio better than 2. Consider such an allocation rule $\mathbf{x}$, and the graph $G_1^{\mathbf{x}}$. This graph has one cycle, with one edge from $s_1 = 0$ to $s_1 = 1$ and one edge from $s_1 = 1$ to $s_1 = 0$. The weight of this cycle is

$$(v_{1a}(0) - v_{1b}(0)) + (v_{1b}(1) - v_{1a}(1)) = (1 - 0) + (H - (H + 1 + \epsilon)) = -\epsilon < 0.$$

Based on Theorem 35, this implies that this allocation rule is not implementable □

Lemma 23. *There exists a setting with two items and two agents with unit-demand and single crossing valuations, such that no randomized truthful mechanism achieves more than* $\frac{\sqrt{2}+2}{4}$ *of the optimal welfare.*

Proof. Consider the setting depicted in Figure C.2, with two agents, 1 and 2, and

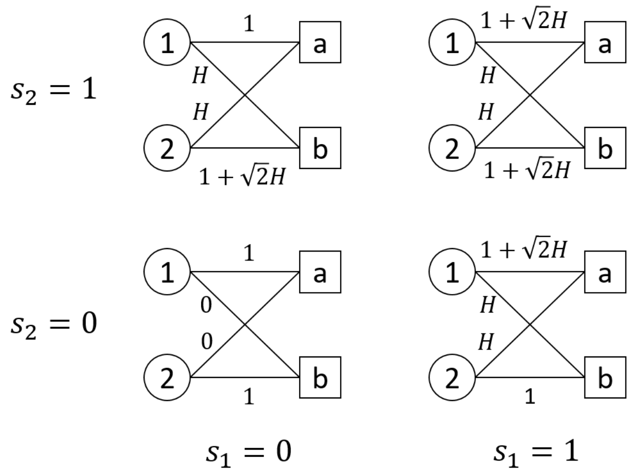


Figure C.2: An instance with unit-demand single-crossing valuations where no randomized truthful allocation achieves more than $\frac{\sqrt{2}+2}{4}$ of the optimal welfare.

two items, a and b. $s_1 \in \{0, 1\}$ and $s_2 \in \{0, 1\}$. The values are

$$\begin{aligned} v_{1a}(0,0) &= 1, & v_{1b}(0,0) &= 0, & v_{2a}(0,0) &= 0, & v_{1b}(0,0) &= 1, \\ v_{1a}(1,0) &= 1+\sqrt{2}H, & v_{1b}(1,0) &= H, & v_{2a}(1,0) &= H, & v_{1b}(1,0) &= 1, \\ v_{1a}(0,1) &= 1, & v_{1b}(0,1) &= H, & v_{2a}(0,1) &= H, & v_{2b}(0,1) &= 1+\sqrt{2}H, \\ v_{1a}(1,1) &= 1+\sqrt{2}H, & v_{1b}(1,1) &= H, & v_{2a}(1,1) &= H, & v_{2b}(1,1) &= 1+\sqrt{2}H, \end{aligned}$$

for an arbitrarily large H. One can easily verify that the valuations are single crossing. We claim that the following equalities hold with respect to the allocation rule of the optimal randomized mechanism:

(a) For every s_1, s_2, $x_{1a}(s_1, s_2) = x_{2b}(s_2, s_1)$ and $x_{2a}(s_1, s_2) = x_{1b}(s_2, s_1)$.

(b) For some $q \in [0, 1]$, $x_{1a}(0, 0) = x_{2b}(0, 0) = q$ and $x_{1\emptyset}(0, 0) = x_{2\emptyset}(0, 0) = 1 - q$.

(c) For some $p \in [0, 1]$, $x_{1a}(0, 1) = p$ and $x_{1b}(0, 1) = 1 - p$.

We next prove the above equalities.

(a) Consider some implementable allocation rule $\bar{x}$, and consider the allocation rule $\tilde{x}$ where $\tilde{x}_{1a}(s_1, s_2) = \bar{x}_{2b}(s_2, s_1)$ and $\tilde{x}_{2a}(s_1, s_2) = \bar{x}_{1b}(s_2, s_1)$ for every s_1, s_2. Note that the valuations are symmetric; i.e., the role of item a (resp. b) for agent 1 is the same as the role of items b (resp. a) for agent 2. By symmetry, $\bar{x}$

is implementable if and only if $\tilde{x}$ is implementable, and both allocation rules have the same approximation guarantee. Clearly, an allocation rule x that applies allocation rules $\bar{x}$ and $\tilde{x}$, with probability $\frac{1}{2}$ each, maintains the same approximation guarantee. Moreover, this allocation rule satisfies the desired property.

(b) The optimal mechanism gains nothing from assigning any positive probability for allocating item b to agent 1 under signal profile $(0,0)$. This is because item b grants no value to agent 1, and in terms of incentives, it can only incentivize agent 1 to misreport his signal at signal profile $(1,0)$. Analogously, the optimal mechanism gains nothing from assigning any positive probability for allocating item a to agent 2 under signal profile $(0,0)$. By (a), $x_{1a}(0,0) = x_{2b}(0,0) = q$ for some $q \in [0,1]$. To conclude the proof of (b), note that the only other feasible set for the agents is the empty set (otherwise, agent 1 has some probability to get item b and agent 2 has some probability to get item a).

(c) Consider $G_1^{\mathbf{x}}$ and the cycle $C = (0,0) \to (1,0) \to (0,0)$ in $G_1^{\mathbf{x}}$. This is the only cycle that contains the node $(1,0)$ in $G_1^{\mathbf{x}}$. Assume $x_{1\emptyset}(1,0) > 0$. Transferring $z \in (0,1]$ probability from $x_{1\emptyset}(1,0)$ to $x_{1a}(1,0)$ decreases the weight of the edge $(0,0) \to (1,0)$ by z, and increases the weight of the edge $(1,0) \to (0,0)$ by $z(1 + \sqrt{2}H) > z$. Therefore, its net effect on the weight of C is positive. Transferring $z \in (0,1]$ probability from $x_{1\emptyset}(1,0)$ to $x_{1b}(1,0)$ does not affect the weight of the edge $(0,0) \to (1,0)$, and increases the weight of the edge $(1,0) \to (0,0)$ by zH. Therefore, its net effect on the weight of C is positive. Since transferring $x_{1\emptyset}(1,0)$ to $x_{1a}(1,0)$ and $x_{1b}(1,0)$ increases welfare and does not violate cycle monotonicity, the optimal mechanism clearly assigns no probability to $x_{1\emptyset}(1,0)$.

Now assume $x_{1\{a,b\}}(1,0) > 0$. By Moving this probability to $x_{1a}(1,0)$, we get the same expected welfare at $(1,0)$, and the weight of the edges in C does not change. Therefore, we may also assume the mechanism does not assign positive utility to $x_{1\{a,b\}}(1,0)$.

According to Theorem 35, in any truthful mechanism, the weight of the cycle C

must be non-negative . This translates to the following condition.

$$\begin{aligned}
& \left(\mathbb{E}_{T\sim x_1(0,0)}[v_{1T}(0,0)] - \mathbb{E}_{T\sim x_1(1,0)}[v_{1T}(0,0)]\right) - \left(\mathbb{E}_{T\sim x_1(1,0)}[v_{1T}(1,0)] - \mathbb{E}_{T\sim x_1(0,0)}[v_{1T}(1,0)]\right) \\
= \quad & (q-p) \;+\; \left(p(1+\sqrt{2}H) + (1-p)H - q(1+\sqrt{2}H)\right) \geq 0 \\
\Rightarrow \quad & q \leq p\left(1 - \frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}.
\end{aligned}$$

In the optimal mechanism, q will be as large as possible in order to maximize the expected welfare at signal profile $(0,0)$. Hence, we can assume $q = p\left(1 - \frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}$. Therefore, the approximation ratio at profile $(0,0)$ is at most $q = p\left(1 - \frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}$. At profile $(0,1)$, if item a is allocated to agent 1 (which happens with probability p), the welfare of the mechanism is at most $2 + \sqrt{2}H$, while the welfare of the optimal allocation is $2H$. As H can be arbitrarily large, this approximation ratio tends to $\frac{1}{\sqrt{2}}$. Therefore, the approximation ratio at profile $(1,0)$ is at most $\frac{p}{\sqrt{2}} + (1-p) = 1 - p\left(1 - \frac{1}{\sqrt{2}}\right)$. The optimal mechanism would balance between the approximation ratio at $(0,0)$ and at $(1,0)$, therefore uses p that solves

$$p\left(1 - \frac{1}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}} = 1 - p\left(1 - \frac{1}{\sqrt{2}}\right).$$

Solving for p, we get $p = \frac{1}{2}$. This leads to an approximation ratio of at most $\frac{2+\sqrt{2}}{4}$, as promised. □

C.2 $n-1$ Lower Bound for Deterministic Mechanisms with Single-Crossing SOS Valuations.

We show that for downward-closed environments, even if valuations satisfy a single-crossing condition and are SOS, any deterministic mechanism cannot obtain a better approximation to the optimal welfare than $n-1$.

Theorem 36. *There exists a downward-closed environment with valuations that satisfy single-crossing for which no deterministic mechanism more than a $n-1$ fraction of the optimal welfare.*

Proof. Consider a set of n bidders, where $\mathcal{I} = \{1\} \cup P(\{2,\ldots,n\})$, where $P(\{2,\ldots,n\})$

is the power set of the set $\{2, \ldots, n\}$. Only agent 1 has a signal $s_1 \in \{0, 1\}$, and other players do not have signals. The valuations are:

$$\begin{aligned} v_1(0) &= 1 \qquad & v_1(1) &= 1 + H \\ v_i(0) &= 0 \qquad & v_i(1) &= H \qquad & \forall i \in \{2, \ldots, n\} \end{aligned}$$

for an arbitrary large value $H \gg 1$. Once can easily verify these valuations satisfy single-crossing and SOS.

Any deterministic mechanism that wants to get any approximation to the social welfare must allocate to agent 1 when $s_1 = 0$. In addition, if a deterministic mechanism wants to get a better approximation than $n - 1$ to the optimal social welfare, agent 1 cannot be allocated when $s_1 = 1$. Otherwise, none of the bidders in $\{2, \ldots n\}$ can get allocated because the only set in $\mathcal{I}$ that contains agent 1 is the singleton set. Therefore, if agent 1 is allocated at $s_1 = 1$, the achieved welfare is $1 + H$, whereas the optimal welfare is $(n - 1) \cdot H$ (when serving all agents in $\{2, \ldots, n\}$). For an arbitrary large H This ratio approaches $n - 1$.

The proof follows since serving agent 1 at $s_1 = 0$ and not serving agent 1 at $s_1 = 1$ is violates monotonicity. □

Remark 37. *The $n - 1$ factor is tight for single-crossing valuations. If $[n] \in \mathcal{I}$, then the mechanism an always allocate all agents. Otherwise, one can always allocate only to the highest valued agent, which is monotone because of single crossing. Since the largest feasible set is of size at most $n - 1$ in this case, allocating to the highest valued agent yields an approximation ratio of $n - 1$.*

C.3 Results for d-SOS

We now extend the results in Section 6.6 to the case of combinatorial d-SOS and combinatorial d-strong-SOS valuations with single-dimensional signals. We first note that if we consider d-SOS valuations, then Equation (6.6) in the decomposition becomes

$$W^* \leq \sum_i v_{iT_i^*}(\mathbf{s}_{-i}, 0_i) + \sum_{i\,:\,s_i>0} d \cdot \big(v_{iT_i^*}(\mathbf{0}_{-i}, s_i) - v_{iT_i^*}(\mathbf{0})\big)$$

$$\leq \underbrace{\sum_i v_{iT_i^*}(\mathbf{s}_{-i}, 0_i)}_{\text{OTHER}} + \underbrace{\sum_{\ell=1}^{k-1} \sum_{i\,:\,s_i=\ell} d \cdot v_{iT_i^*}(\mathbf{0}_{-i}, s_i)}_{\text{SELF}}, \tag{C.2}$$

We now show the extension of Theorem 30 to d-SOS valuations.

Theorem 38. *For every combinatorial auction with d-SOS valuations over single-dimensional signals, and signal space of size k, i.e., $s_i \in \{0, 1, \ldots, k-1\}\, \forall i$, there exists a truthful mechanism that gives $d(k+1)+2$-approximation to the optimal social welfare.*

Proof. The mechanism is identical to k-HL, but runs (`Random Threshold`) with probability $p_{RT} = \frac{(k-1)d}{d(k+1)+2}$ and (`Random Sampling`) With probability $1 - p_{RT}$. The mechanism was already proved to be truthful in Section 6.6.1.

`Random Threshold` now gives a $d(k-1)$-approximation to the new SELF term. The proof is the same as of Lemma 19, but the extra factor of d comes from the fact the the new SELF term is d times larger.

`Random Sampling` gives a $2(d+1)$-approximation to the OTHER term. While this term is the same for d-SOS, the new factor is due to the fact that when applying Lemma 18 in the proof of Lemma 20, we get that $\mathbb{E}_{A,B}[\tilde{v}_{iT}] \geq \frac{1}{2(d+1)} v_{iT}(\mathbf{s}_{-i}, 0_i)$ instead of the bound we get in Equation (6.8).

The new approximation guarantee follows from the new decomposition, the new approximation guarantees the various mechanisms get for the terms of the decomposition, and the updated probability p_{RT}. □

We next extend Theorem 31.

Theorem 39. *For every combinatorial auction with d-strong-SOS valuations over single-dimensional signals, and signal space of size k, i.e., $s_i \in \{0, 1, \ldots, k-1\}\, \forall i$, there exists a truthful mechanism that gives $(d(d+1)\log_2 k + 2(d+1))$-approximation to the optimal social welfare.*

Proof. The mechanism is identical to mechanism k-SS from Section 6.6.2, but runs `Random Bucket` with probability $p_{RB} = \frac{d \log_2 k}{d \log_2 k + 2}$ and (`Random Sampling`) With probability $1 - p_{RB}$.

The SELF term from Equation (6.9) is now bounded via the following:

$$\begin{aligned} \mathsf{SELF} &= \sum_{\ell=1}^{k-1} \sum_{i\,:\,s_i=\ell} d \cdot v_{iT_i^*}(\mathbf{0}_{-i}, s_i) \\ &= \sum_{\ell=1}^{\log_2 k} \sum_{i\,:\,2^{\ell-1}\le s_i<2^\ell} d \cdot v_{iT_i^*}(\mathbf{0}_{-i}, s_i) \\ &\le \sum_{\ell=1}^{\log_2 k} \sum_{i\,:\,2^{\ell-1}\le s_i<2^\ell} d(d+1) \cdot v_{iT_i^*}(\mathbf{0}_{-i}, {2^{\ell-1}}_i), \end{aligned} \tag{C.3}$$

where the inequality follows the definition of d-strong-SOS valuations.

The new bound changes the guarantee of `Random Bucket` to get a $d(d+1)\log_2 k$-approximation to the SELF term, where the proof is identical to that of Lemma 21.

As stated in Theorem 38, `Random Sampling` approximates the OTHER term to a factor $2(d+1)$. The proof of the new bound follows the new decomposition, the updated probabilities and the new approximation guarantees of the mechanisms being run. □

Bibliography

Ittai Abraham, Susan Athey, Moshe Babaioff, and M Grubb. Peaches. *Lemons, and Cookies: Designing Auction Markets with Dispersed Information*, 2011.

Ittai Abraham, Moshe Babaioff, Shaddin Dughmi, and Tim Roughgarden. Combinatorial auctions with restricted complements. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 3–16. ACM, 2012.

S. Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 512–521, Oct 2011. doi: 10.1109/FOCS.2011.90.

Edward J Anderson and Peter Nash. *Linear programming in infinite-dimensional spaces: theory and applications*. John Wiley & Sons, 1987.

Mark Armstrong. Multiproduct Nonlinear Pricing. *Econometrica*, 64(1):51–75, 1996. ISSN 00129682. doi: 10.2307/2171924.

Susan Athey. Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69(4):861–889, 2001. ISSN 00129682, 14680262.

Lawrence M Ausubel et al. A generalized vickrey auction. *Econometrica*, 1999.

Moshe Babaioff, Robert Kleinberg, and Renato Paes Leme. Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 92–109, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1415-2.

Moshe Babaioff, Nicole Immorlica, Brendan Lucier, and S. Matthew Weinberg. A Simple and Approximately Optimal Mechanism for an Additive Buyer. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 21–30. IEEE, 2014.

Moshe Babaioff, Yannai A. Gonczarowski, and Noam Nisan. The menu-size complexity of revenue approximation. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 869–877, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4528-6. doi: 10.1145/3055399.3055426. URL `http://doi.acm.org/10.1145/3055399.3055426`.

Dirk Bergemann and Stephen Morris. Robust mechanism design. *Econometrica*, 73(6):1771–1813, 2005.

Dirk Bergemann, Xianwen Shi, and Juuso Välimäki. Information acquisition in interdependent value auctions. *Journal of the European Economic Association*, 7(1):61–89, 2009.

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. The Limits of Price Discrimination âĂă. *American Economic Review*, 105(3):921–957, mar 2015. ISSN 0002-8282. doi: 10.1257/aer.20130848. URL https://www.aeaweb.org/articles.php?doi=10.1257/aer.20130848.

Andrew An Bian, Kfir Yehuda Levy, Andreas Krause, and Joachim M. Buhmann. Non-monotone continuous dr-submodular maximization: Structure and algorithms. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pages 486–496, 2017. URL http://papers.nips.cc/paper/6652-non-monotone-continuous-dr-submodular-maximization-structure-and-algorithm

Sushil Bikhchandani. Ex post implementation in environments with private goods. 2006.

Patrick Briest, Shuchi Chawla, Robert Kleinberg, and S. Matthew Weinberg. Pricing lotteries. *Journal of Economic Theory*, 156:144–174, 2015.

Johannes Brustle, Yang Cai, Fa Wu, and Mingfei Zhao. Approximating gains from trade in two-sided markets via simple mechanisms. In *Proceedings of the 2017 ACM Conference on Economics and Computation, EC '17, Cambridge, MA, USA, June 26-30, 2017*, pages 589–590, 2017. doi: 10.1145/3033274.3085148. URL http://doi.acm.org/10.1145/3033274.3085148.

Jeremy Bulow and John Roberts. The simple economics of optimal auctions. *Journal of political economy*, 97(5):1060–1090, 1989.

Yang Cai and C. Daskalakis. Extreme-value theorems for optimal multidimensional pricing. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 522–531, Oct 2011. doi: 10.1109/FOCS.2011.76.

Yang Cai and Zhiyi Huang. Simple and nearly optimal multi-item auctions. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '13, pages 564–577, 2013. URL http://epubs.siam.org/doi/abs/10.1137/1.9781611973105.41.

Yang Cai and Mingfei Zhao. Simple Mechanisms for Subadditive Buyers via Duality. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 170–183, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4528-6. doi: 10.1145/3055399.3055465. URL http://doi.acm.org/10.1145/3055399.3055465.

Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. On optimal multidimensional mechanism design. *SIGecom Exch.*, 10(2):29–33, June 2011. ISSN 1551-9031. doi: 10.1145/1998549.1998555. URL http://doi.acm.org/10.1145/1998549.1998555.

Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Optimal multidimensional mechanism design: Reducing revenue to welfare maximization. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 130–139, 2012. doi: 10.1109/FOCS.2012.88. URL http://dx.doi.org/10.1109/FOCS.2012.88.

Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Understanding incentives: Mechanism design becomes algorithm design. *CoRR*, abs/1305.4002, 2013a. URL http://arxiv.org/abs/1305.4002.

Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Reducing revenue to welfare maximization: Approximation algorithms and other generalizations. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '13, pages 578–595. SIAM, 2013b. ISBN 978-1-611972-51-1. URL http://dl.acm.org/citation.cfm?id=2627817.2627859.

Yang Cai, Nikhil R. Devanur, and S. Matthew Weinberg. A Duality Based Unified Approach to Bayesian Mechanism Design. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 926–939, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4132-5. doi: 10.1145/2897518.2897645. URL http://doi.acm.org/10.1145/2897518.2897645.

Yang Cai, Nikhil R. Devanur, Kira Goldner, and R. Preston McAfee. Simple and approximately optimal pricing for proportional complementarities. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, EC '19, Phoenix, AZ, USA, 2019. ACM. ISBN 978-1-4503-6792-9. doi: 10.1145/3328526.3329562. URL https://homes.cs.washington.edu/~kgoldner/papers/proportionalcomplements.pdf.

Shuchi Chawla and J. Benjamin Miller. Mechanism design for subadditive agents via an ex ante relaxation. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, pages 579–596, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-3936-0. doi: 10.1145/2940716.2940756. URL http://doi.acm.org/10.1145/2940716.2940756.

Shuchi Chawla and Balasubramanian Sivan. Bayesian algorithmic mechanism design. *SIGecom Exch.*, 13(1):5–49, November 2014. ISSN 1551-9031. doi: 10.1145/2692375.2692378. URL `http://doi.acm.org/10.1145/2692375.2692378`.

Shuchi Chawla, Jason D. Hartline, and Robert Kleinberg. Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, EC '07, pages 243–251, 2007. URL `http://doi.acm.org/10.1145/1250910.1250946`.

Shuchi Chawla, Jason D Hartline, David L Malec, and Balasubramanian Sivan. Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 311–320. ACM, 2010. URL `http://dl.acm.org/citation.cfm?id=1806733`.

Shuchi Chawla, Hu Fu, and Anna Karlin. Approximate revenue maximization in interdependent value settings. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, pages 277–294, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2565-3.

Shuchi Chawla, David Malec, and Balasubramanian Sivan. The power of randomness in bayesian optimal mechanism design. *Games and Economic Behavior*, 91:297 – 317, 2015. ISSN 0899-8256. doi: http://dx.doi.org/10.1016/j.geb.2012.08.010. URL `http://www.sciencedirect.com/science/article/pii/S0899825612001303`.

Yeon-Koo Che and Ian Gale. The optimal mechanism for selling to a budget-constrained buyer. *Journal of Economic theory*, 92(2):198–233, 2000.

Yeon-Koo Che, Jinwoo Kim, and Fuhito Kojima. Efficient assignment with interdependent values. *Journal of Economic Theory*, 158:54–86, 2015.

Edward H Clarke. Multipart pricing of public goods. *Public choice*, 11(1):17–33, 1971.

Vincent Conitzer, Tuomas Sandholm, and Paolo Santi. Combinatorial auctions with k-wise dependent valuations. In *AAAI*, volume 5, pages 248–254, 2005.

Florin Constantin and David C. Parkes. On revenue-optimal dynamic auctions for bidders with interdependent values. In *AMEC/TADA*, volume 13 of *Lecture Notes in Business Information Processing*, pages 1–15. Springer, 2007.

Florin Constantin, Takayuki Ito, and David C. Parkes. Online auctions for bidders with interdependent values. In *AAMAS*, page 110. IFAAMAS, 2007.

Jacques Crémer and Richard P. McLean. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345–361, 1985. ISSN 00129682, 14680262.

Jacques Crémer and Richard P. McLean. Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257, 1988. ISSN 00129682, 14680262.

Partha Dasgupta and Eric Maskin. Efficient auctions. *The Quarterly Journal of Economics*, 115(2):341–388, 2000.

Constantinos Daskalakis. Multi-item auctions defying intuition? *ACM SIGecom Exchanges*, 14(1):41–75, 2015.

Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. Mechanism design via optimal transport. In *Proceedings of the 14th ACM Conference on Electronic Commerce*, EC '13, pages 269–286, 2013. URL `http://doi.acm.org/10.1145/2482540.2482593`.

Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. Strong duality for a multiple-good monopolist. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 449–450. ACM, 2015.

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

C d'Aspremont and L.-A Gérard-Varet. Bayesian incentive compatible beliefs. *Journal of Mathematical Economics*, 10(1):83 – 103, 1982. ISSN 0304-4068.

Nikhil R. Devanur and S. Matthew Weinberg. The Optimal Mechanism for Selling to a Budget Constrained Buyer: The General Case. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 39–40, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4527-9. doi: 10.1145/3033274.3085132. URL `http://doi.acm.org/10.1145/3033274.3085132`.

Nikhil R. Devanur, Nima Haghpanah, and Christos-Alexandros Psomas. Optimal Multi-Unit Mechanisms with Private Demands. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 41–42, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4527-9. doi: 10.1145/3033274.3085122. URL `http://doi.acm.org/10.1145/3033274.3085122`.

Nikhil R. Devanur, Kira Goldner, Raghuvansh R. Saxena, Ariel Schvartzman, and S. Matthew Weinberg. Optimal mechanism design for single-minded agents. In *Manuscript*, 2019. URL `https://homes.cs.washington.edu/~kgoldner/papers/partiallyordered.pdf`.

Shahar Dobzinski, Hu Fu, and Robert D. Kleinberg. Optimal auctions with correlated bidders are easy. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, pages 129–138, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0691-1.

Alon Eden, Michal Feldman, Ophir Friedler, Inbal Talgam-Cohen, and S. Matthew Weinberg. The Competition Complexity of Auctions: A Bulow-Klemperer Result for Multi-Dimensional Bidders. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 343–343, New York, NY, USA, 2017a. ACM. ISBN 978-1-4503-4527-9. doi: 10.1145/3033274.3085115. URL http://doi.acm.org/10.1145/3033274.3085115.

Alon Eden, Michal Feldman, Ophir Friedler, Inbal Talgam-Cohen, and S. Matthew Weinberg. A Simple and Approximately Optimal Mechanism for a Buyer with Complements: Abstract. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, EC '17, pages 323–323, New York, NY, USA, 2017b. ACM. ISBN 978-1-4503-4527-9. doi: 10.1145/3033274.3085116. URL http://doi.acm.org/10.1145/3033274.3085116.

Alon Eden, Michal Feldman, Amos Fiat, and Kira Goldner. Interdependent values without single-crossing. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, EC '18, pages 369–369, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5829-3. doi: 10.1145/3219166.3219173. URL https://arxiv.org/pdf/1806.03865.pdf.

Alon Eden, Michal Feldman, Amos Fiat, Kira Goldner, and Anna R. Karlin. Combinatorial auctions with interdependent valuations: Sos to the rescue. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, EC '19, Phoenix, AZ, USA, 2019. ACM. ISBN 978-1-4503-6792-9. doi: 10.1145/3328526.3329759. URL https://arxiv.org/abs/1903.08384.

Uriel Feige. On maximizing welfare when utility functions are subadditive. *SIAM Journal on Computing*, 39(1):122–142, 2009.

Uriel Feige and Rani Izsak. Welfare maximization and the supermodular degree. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 247–256. ACM, 2013.

Uriel Feige, Michal Feldman, Nicole Immorlica, Rani Izsak, Brendan Lucier, and Vasilis Syrgkanis. A unifying hierarchy of valuations with complements and substitutes. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

Amos Fiat, Kira Goldner, Anna R. Karlin, and Elias Koutsoupias. The FedEx Problem. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, pages 21–22, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-3936-0. doi: 10.1145/2940716.2940752. URL `http://doi.acm.org/10.1145/2940716.2940752`.

Hu Fu, Chris Liaw, Pinyan Lu, and Zhihao Gavin Tang. The value of information concealment. *arXiv preprint arXiv:1707.05875*, 2017.

Yiannis Giannakopoulos and Elias Koutsoupias. Duality and Optimality of Auctions for Uniform Distributions. In *Proceedings of the 15th ACM Conference on Economics and Computation*, EC '14, pages 259–276, 2014. URL `http://doi.acm.org/10.1145/2600057.2602883`.

Yiannis Giannakopoulos and Elias Koutsoupias. Selling Two Goods Optimally. In *ICALP (2)*, volume 9135 of *Lecture Notes in Computer Science*, pages 650–662. Springer, 2015.

Andrew V. Goldberg, Jason D. Hartline, and Andrew Wright. Competitive auctions and digital goods. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 735–744, Philadelphia, PA, USA, 2001. ISBN 0-89871-490-7.

Yannai A. Gonczarowski. Bounding the menu-size of approximately optimal auctions via optimal-transport duality. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 123–131, 2018. doi: 10.1145/3188745.3188786. URL `http://doi.acm.org/10.1145/3188745.3188786`.

Theodore Groves. Incentives in teams. *Econometrica: Journal of the Econometric Society*, pages 617–631, 1973.

Nima Haghpanah and Jason Hartline. Reverse mechanism design. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 757–758. ACM, 2015. Updated version: http://arxiv.org/abs/1404.1341.

Sergiu Hart and Noam Nisan. The menu-size complexity of auctions. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 565–566, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1962-1. doi: 10.1145/2482540.2482544. URL `http://doi.acm.org/10.1145/2482540.2482544`.

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

Sergiu Hart and Philip J. Reny. Maximal revenue with multiple goods: Non-monotonicity and other observations. 2012. URL http://www.ma.huji.ac.il/hart/abs/monot-m.html.

Jason D. Hartline. Mechanism design and approximation. *Book draft. October*, 122, 2013. URL http://jasonhartline.com/MDnA/.

Takayuki Ito and David C. Parkes. Instantiating the contingent bids model of truthful interdependent value auctions. In *AAMAS*, pages 1151–1158. ACM, 2006.

Philippe Jehiel and Benny Moldovanu. Efficient design with interdependent valuations. *Econometrica*, 69(5):1237–1259, 2001.

Philippe Jehiel, Moritz Meyer-ter Vehn, Benny Moldovanu, and William R Zame. The limits of ex post implementation. *Econometrica*, 74(3):585–610, 2006.

Brian Kelly. Why Your Uber Driver May Force You to Cancel Your Ride and Pay a Fee. https://thepointsguy.com/news/uber-drivers-cancellation-fee-scam/, 2018.

David Kempe, Vasilis Syrgkanis, and Eva Tardos. Information asymmetries in common-value auctions with discrete signals. *SSRN eLibrary*, 2013.

Mark Klein, Gabriel A. Moreno, David C. Parkes, Daniel Plakosh, Sven Seuken, and Kurt C. Wallnau. Handling interdependent values in an auction mechanism for bandwidth allocation in tactical data networks. In *NetEcon*, pages 73–78. ACM, 2008.

Paul Klemperer. Auctions with almost common values: The wallet game'and its applications. *European Economic Review*, 42(3):757–769, 1998.

Vijay Krishna. *Auction theory*. Academic press, 2009.

Jean-Jacques Laffont, Eric Maskin, and Jean-Charles Rochet. Optimal nonlinear pricing with two-dimensional characteristics. *Information, Incentives and Economic Mechanisms*, pages 256–266, 1987.

Daniel Lehmann, Liadan Ita O'Callaghan, and Yoav Shoham. Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM (JACM)*, 49(5):577–602, 2002.

Xinye Li and Andrew Chi-Chih Yao. On revenue maximization for selling multiple independently distributed items. *Proceedings of the National Academy of Sciences*, 110(28):11232–11237, 2013. URL http://dx.doi.org/10.1073/pnas.1309533110.

Yunan Li. Approximation in mechanism design with interdependent values. *Games and Economic Behavior*, 2016.

Siqi Liu and Christos-Alexandros Psomas. On the competition complexity of dynamic mechanism design on the competition complexity of dynamic mechanism design on the competition complexity of dynamic mechanism design on the competition complexity of dynamic mechanism design. *arXiv preprint arXiv:1709.07955*, 2017.

David G Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1997.

Will Ma and David Simchi-Levi. Reaping the benefits of bundling under high production costs. *arXiv preprint arXiv:1512.02300*, 2015.

Alexey Malakhov and Rakesh V Vohra. An optimal auction for capacity constrained bidders: a network perspective. *Economic Theory*, 39(1):113–128, 2009.

Alejandro M. Manelli and Daniel R. Vincent. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127 (1):1 – 35, 2006. URL `http://dx.doi.org/10.1016/j.jet.2005.08.007`.

Alejandro M. Manelli and Daniel R. Vincent. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory*, 137(1):153 – 185, 2007. URL `http://dx.doi.org/10.1016/j.jet.2006.12.007`.

Eric Maskin. Auctions and privatization. *Privatization, H. Siebert, ed. (Institut fur Weltwirtschaften der UniversitaÂÎt Kiel: 1992)*, pages 115–136, 1992.

R. Preston McAfee and John McMillan. Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, 46(2):335 – 354, 1988. URL `http://dx.doi.org/10.1016/0022-0531(88)90135-4`.

Richard McLean and Andrew Postlewaite. Implementation with interdependent valuations. *Theoretical Economics*, 10(3), 2015.

Paul Milgrom. *Putting auction theory to work*. Cambridge University Press, 2004.

Paul R Milgrom and Robert J Weber. A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122, 1982.

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. URL `http://dx.doi.org/10.1287/moor.6.1.58`.

Rad Niazadeh, Tim Roughgarden, and Joshua R. Wang. Optimal algorithms for continuous non-monotone submodular and dr-submodular maximization. In *Annual Conference on Neural Information Processing Systems 2018, 3-8 December 2018, Montréal, Canada.*, pages 9617–9627, 2018.

Christos H. Papadimitriou and George Pierrakos. On optimal single-item auctions. In *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC 2011, San Jose, CA, USA, 6-8 June 2011*, pages 119–128, 2011.

Gregory Pavlov. Optimal mechanism for selling two goods. *The B.E. Journal of Theoretical Economics*, 11(3), 2011. URL `http://dx.doi.org/10.2202/1935-1704.1664`.

Marek Pycia. *Stochastic vs deterministic mechanisms in multidimensional screening*. PhD thesis, MIT, 2006. URL `http://pycia.bol.ucla.edu/pycia-multidimensional-screening.pdf`. Chapter 3.

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

Valentin Robu, David C. Parkes, Takayuki Ito, and Nicholas R. Jennings. Efficient interdependent value combinatorial auctions with single minded bidders. In *IJCAI*, pages 339–345. IJCAI/AAAI, 2013.

Jean-Charles Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of mathematical Economics*, 16(2):191–200, 1987.

R Tyrrell Rockafellar. *Conjugate duality and optimization*, volume 16. Siam, 1974.

Amir Ronen. On approximating optimal auctions. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 11–17. ACM, 2001.

Tim Roughgarden. Approximately optimal mechanism design: Motivation, examples, and lessons learned. *SIGecom Exch.*, 13(2):4–20, January 2015. ISSN 1551-9031. doi: 10.1145/2728732.2728733. URL `http://doi.acm.org/10.1145/2728732.2728733`.

Tim Roughgarden and Inbal Talgam-Cohen. Optimal and robust mechanism design with interdependent values. *ACM Trans. Econ. Comput.*, 4(3):18:1–18:34, June 2016. ISSN 2167-8375.

Aviad Rubinstein and S. Matthew Weinberg. Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 377–394. ACM, 2015.

Raghuvansh R. Saxena, Ariel Schvartzman, and S. Matthew Weinberg. The menu-complexity of one-and-a-half dimensional mechanism design. In *To appear in Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '18, 2018.

Tasuku Soma and Yuichi Yoshida. A generalization of submodular cover via the diminishing return property on the integer lattice. In *Advances in Neural Information Processing Systems 28*, pages 847–855. Curran Associates, Inc., 2015.

John Thanassoulis. Haggling over substitutes. *Journal of Economic Theory*, 117 (2):217 – 245, 2004. URL http://www.sciencedirect.com/science/article/pii/S0022053103003351.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961.

Rakesh Vohra. Paths, cycles and mechanism design. *Preprint*, 2007.

Zihe Wang and Pingzhong Tang. Optimal mechanisms with simple menus. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC '14, pages 227–240, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2565-3. doi: 10.1145/2600057.2602863. URL http://doi.acm.org/10.1145/2600057.2602863.

Robert B Wilson. Competitive Bidding with Disparate Information. *Management Science*, 15(7):446–452, 1969.

Andrew Chi-Chih Yao. An n-to-1 Bidder Reduction for Multi-item Auctions and its Applications. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 92–109. SIAM, 2015. doi: 10.1137/1.9781611973730.8. URL http://arxiv.org/abs/1406.3278.