Optimal Mechanism Design for Single-Minded Agents

NIKHIL R. DEVANUR*, Amazon
KIRA GOLDNER†, Columbia University
RAGHUVANSH R. SAXENA‡, Princeton University
ARIEL SCHVARTZMAN§, Princeton University
S. MATTHEW WEINBERG¶, Princeton University

We consider optimal (revenue maximizing) mechanism design in the interdimensional setting, where one dimension is the 'value' of the buyer, and the other is a 'type' that captures some auxiliary information. A prototypical example of this is the FedEx Problem, for which Fiat et al. [10] characterize the optimal mechanism for a single agent. Another example of this is when the type encodes the buyer's budget [DW17]. The question we address is *how far can such characterizations go?* In particular, we consider the setting of single-minded agents. A seller has heterogenous items. A buyer has a valuation v for a specific subset of items S, and obtains value v if and only if he gets all the items in S (and potentially some others too).

We show the following results.

(1) Deterministic mechanisms (i.e. posted prices) are optimal for distributions that satisfy the "declining marginal revenue" (DMR) property. In this case we give an explicit construction of the optimal mechanism.
(2) Without the DMR assumption, the result depends on the structure of the minimal directed acyclic graph (DAG) representing the partial order among types. When the DAG has out-degree at most 1, we characterize the optimal mechanism à la FedEx; this can be thought of as a generalization of the FedEx characterization since FedEx corresponds to a DAG that is a line.
(3) Surprisingly, without the DMR assumption *and* when the DAG has at least one node with an out-degree of at least 2, then we show that there is no hope of such a characterization. The minimal such example happens on a DAG with 3 types. We show that in this case the menu complexity is *unbounded* in that for any M, there exist distributions over (v, S) pairs such that the menu complexity of the optimal mechanism is at least M.
(4) For the case of 3 types, we also show that for all distributions there exists an optimal mechanism of *finite* menu complexity. This is in contrast to the case where you have 2 heterogenous items with additive utilities for which the menu complexity could be uncountably infinite [DDT15, MV07].

In addition, we prove that optimal mechanisms for Multi-Unit Pricing (without a DMR assumption) can have unbounded menu complexity as well, and we further propose an extension where the menu complexity of optimal mechanisms can be countably infinite, but not uncountably infinite.

*iam@nikhildevanur.com.
†kgoldner@cs.columbia.edu. Supported in part by NSF CCF-1420381 and by a Microsoft Research PhD Fellowship. Supported in part by NSF award DMS-1903037 and a Columbia Data Science Institute postdoctoral fellowship.
‡rrsaxena@princeton.edu.
§acohenca@princeton.edu. Supported by NSF CCF-1717899.
¶smweinberg@princeton.edu. Supported by NSF CCF-1717899.

EC '20, July 13–17, 2020, Virtual Event, Hungary

ACM ISBN 978-1-4503-7975-5/20/07…$15.00
https://doi.org/10.1145/3391403.3399454

Taken together, these results establish that optimal mechanisms in interdimensional settings are both surprisingly richer than single-dimensional settings, yet also vastly more structured than multi-dimensional settings.

CCS Concepts: • **Theory of computation → Algorithmic mechanism design**.

Additional Key Words and Phrases: optimal mechanism design, revenue, duality, interdimensional, menu complexity, single-minded valuations, partial Lagrangian.

ACM Reference Format:
Nikhil R. Devanur, Kira Goldner, Raghuvansh R. Saxena, Ariel Schvartzman, and S. Matthew Weinberg. 2020. Optimal Mechanism Design for Single-Minded Agents. In *Proceedings of the 21st ACM conference on Economics and Computation (EC '20), July 13–17, 2020, Virtual Event, Hungary*. ACM, New York, NY, USA, 64 pages. https://doi.org/10.1145/3391403.3399454

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 [22] 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 suffer from numerous sources of intractability (including computational intractability, non-monotonicity, high description complexity, and others): [2, 6, 7, 14, 15, 20].

Very recently, Fiat et al. [10] identify a fascinating 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). We can think of the input as being a (correlated) 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 getting internet service. So you value options such as exclusively internet service, internet/phone service, or internet/cable, and so on, at v. For any option that does not include internet you get a value of zero (so we again think of the input distribution as a 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 2 in Appendix G). 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 the paper.

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.

1.1. Main Results

Fiat et al. [10] give a characterization of an optimal mechanism for the FedEx problem, and our goal is to understand the generalizability of this characterization, in particular to the partially ordered setting. Towards this, we first describe the FedEx characterization and what a generalization could look like. A *deterministic* mechanism sets a posted price p_i for each shipping option, and the buyer picks the option he prefers (if any). Clearly, it makes sense for the prices to be non-increasing in i-day shipping. The FedEx solution is recursive: start with the price on day 1 (as a variable), and constrain the price on day 2 to be weakly lower, and so on. When the distributions satisfy the *Declining Marginal Revenue (DMR)*[1] property, this strategy actually results in deterministic prices that are optimal. Without any distributional assumption, one might have to resort to *lotteries*: the buyer gets the item only with some probability. The first day price is still deterministic, but for the second day, the mechanism offers a lottery such that the *expected* price for full service is weakly lower. It turns out that we only need to randomize between two options. Recursively, every option on day i may split into two options on day $i+1$, so we might have at most 2^{m-1} options on day m, and $2^m - 1$ options overall (and examples exist where $2^m - 1$ options are necessary [23]).

So our starting point is a hope that similar recursive ideas can characterize optimal auctions beyond the totally-ordered FedEx setting. Some terminology is useful here to understand precisely what this might mean. We use the directed acyclic graph (DAG) representation of a partial order: an edge from i to j implies j *is preferred over* i. The DAG is *minimal*: if (i,j) and (j,k) are edges then (i,k) is *not* an edge. The DAG for the FedEx problem goes *right to left*, i.e., it has edges $(i+1,i)$ for i from 1 to $m-1$. A recursive approach for a DAG would look like this: start with a *sink*, set a deterministic price, and use this to constrain the prices (either deterministically or in expectation, based on the distributional assumption) for its predecessors and so on. The goal of this paper is to understand *Will something like this work for partially-ordered items?*

DMR: Under the DMR assumption, this strategy for pricing works in any DAG (Theorem C.1 in Appendix C). We start from the sink nodes and recursively constrain the price of a node to be at most the minimum among the prices of all its successors. Our proof that this procedure works employs LP duality, and a significantly more involved procedure to set appropriate dual variables than in [FGKK16]. The fact that optimal mechanisms are deterministic subject to DMR matches prior work for totally-ordered settings [Che and Gale 2000, DHP17, DW17, FGKK16].

Out-degree 1: The FedEx strategy still works when the minimal DAG has out-degree at most 1, without any distributional assumptions (Theorem D.1 in Appendix D). Compared to the DMR case, we now have to deal with lotteries but when we process a node, there is exactly one successor that constrains the lotteries for this node, in exactly the same way as in FedEx.

3 node DAG: The minimal example where the out-degree is 2 is a three-node DAG with nodes A, B, and C, and edges (C,A) and (C,B). One might hope that the following recursive strategy would work (after all, the graph is still a DAG, and should be amenable to recursive arguments): set prices

[1] A one-dimensional distribution F satisfies Declining Marginal Revenues if $v(1-F(v))$ is concave. Note that this is an *incomparable* condition to Regularity (in the Myersonian sense), and also Monotone Hazard Rate (MHR). There exist distributions which are DMR, but not Regular (and therefore not MHR). There also exist distributions which are MHR but not DMR. See [8] 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	Coordinated, 3 Items	Additive
Det. under DMR	N/A	✓	✓	✓	✗	N/A
Lower Bound	1	$2^m - 1$	**unbounded**	**unbounded**	**countably infinite**	uncountable
Upper Bound	1	$2^m - 1$	**finite**	—	**countably infinite**	uncountable

Bold results are from this paper.

deterministically for A and B and use the minimum of the two to constrain the expected price for C. Note that if there were no item B, this would match precisely the FedEx solution.

It turns out that this idea fails horribly, for the following (very high-level) reason. With just two items (C and A), the price of A transparently constrains what prices we can set for C (the expected price for C must be lower). So when optimizing the price of A, we can take this into account. With three items, it's no longer clear how the price of A constrains the price of C. Certainly, the expected price for C must be lower, but perhaps a stronger constraint is already implied by the price of B. Therefore, one cannot separately optimize the price of A without knowing the price of B.

Indeed, this intuition actually manifests into a lower bound: it is not only challenging to jointly optimize the prices of A, B together, but the optimum may no longer be deterministic at all! Specifically, for any integer M, there exist value distributions for this 3-node DAG for which the unique optimal mechanism presents M different lotteries to the buyer (Theorem 4.1). Essentially, *there is no hope for a FedEx style solution even for this minimal case.* We focus the technical presentation of our paper on this result.

Finiteness of menu complexity: The use of menu complexity lower bounds to ascertain complexity of mechanisms is not new: Daskalakis et al. [7], Manelli and Vincent [20] show that the optimal mechanism for the multi-dimensional setting might have *uncountable* menu complexity—this holds even for just two items with *additive* valuations, 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.

Within the context of these results, we ask if we can get an infinite (uncountable or countable) menu complexity for the partially-ordered setting as well. A natural strategy is to take the limit of our construction as the number of randomizations goes to infinity. Somewhat surprisingly, the example then collapses and has a deterministic price as optimal. We show that this is no coincidence: that the menu complexity for the three item case is *always finite* (Theorem 4.2).

Summary: The main technical takeaway from our results is a thorough understanding of optimal mechanisms in interdimensional settings beyond FedEx 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 more structured than optimal multi-dimensional auctions.

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 H.1). More specifically, given desired properties of a dual λ such as: $\lambda(x) = 0$ for $x \in S$, $\lambda(x) > 0$ for $x \in T$, etc. (see Theorem H.1 for formal statement), the Master Theorem asserts that there exists *some* distribution for which there exists a feasible dual with the desired properties. 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 (reasoning about primals and duals using complementary slackness) from the tedious parts (confirming that there exists an instance for which a dual with particular desired properties even exists).

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 [DHP17] is also unbounded (Theorem I.1 in Appendix I). 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). Second, we propose a generalization beyond totally-ordered settings which we call *coordinated* valuations, and again characterize the menu complexity of optimal mechanisms for one instance of three items (which can be countably infinite, but not uncountable, see Appendix J).

1.3. Related Work

Single-minded valuations are a well-known model (e.g. [18]). Most work in this model pertains to welfare maximization in more complex settings, such as combinatorial auctions. Other work assumes that the buyer's interest is publicly known; in this case, the buyer is single-parameter, and a single-buyer revenue maximization problem reduces to Myerson.

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 [5, 8–10]. The present paper 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 [9] in that (1) both are the only works to use the analysis from [CDW16] 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 [6, 7, 11, 13, 17, 19, 21]. From this

set, the most related are [13, 19], 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 items [1, 2, 7, 12, 14, 24]. Very recent work of [SSW18] considers the menu complexity of approximately optimal mechanisms for the FedEx Problem (for which [FGKK16] 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) [FGKK16], and heterogeneous items (uncountable) [DDT15]. Previously, no settings with this property were known.

1.4. Roadmap

Our paper contains four main results, although we view the primary contributions as (3) and (4):

(1) In Appendix C, we prove Theorem C.1, which explicitly constructs a deterministic optimal auction for partially-ordered items when all marginals are DMR.
(2) In Appendix D, we prove Theorem D.1, which extends the recursive FedEx algorithm for totally-ordered items to partially-ordered items when minimal DAGs with outdegree at most one.
(3) We focus our technical presentation on the ideas necessary for Theorem 4.1, which establishes that any partially-ordered instance for which some node in the minimal DAG has outdegree at least two, the menu complexity of the optimal mechanism may be unbounded. In Section 2 we provide the minimal preliminaries to understand the main ideas behind our proof of this result (full preliminaries in Appendix A). In Section 3 we overview the key duality aspects. In Section 4 we give a brief overview of the proof of Theorem 4.1. The full proof is in Appendix E.
(4) Finally, we also establish that the menu complexity of optimal mechanisms for this minimal 3-item instance is always finite. The main ideas appear in Section 4, and a full proof of Theorem 4.2 appears in Appendix F.

Outside of our main results, Appendix H presents our "Master Theorem" (Theorem H.1), which is of independent interest for future work on mechanism design with totally- or partially-ordered items. In Appendix I and Appendix J we display the applicability of our techniques for related settings such as Multi-Unit Pricing (Theorem I.1) and coordinated values (Theorems J.2, J.4, J.5), respectively. Section 5 presents our conclusions and discusses future work.

2. Preliminaries

In the interest of presentation, we'll provide the minimum preliminaries here for the reader to understand the key ideas. In Appendix A, we provide full preliminaries, including additional intuition, and covering prior work (such as [DW17, FGKK16]). Many of the facts we will use are stated here without proof (proofs are given in Appendix A).

2.1. 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).

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.[2] 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. [10], 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.[3] 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). 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 $a_G(v) = q$.

2.2. Incentive Compatibility, Revenue Curves, and Ironing

As observed in [FGKK16], 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 [FGKK16] 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 $G' \supseteq G$, but there does not exist a G'' with $G' \supseteq G'' \supseteq G$. 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').[4] We conclude this section with some standard definitions and observations.

Definition 2.1 (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*

[2]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.

[3]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. 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.

[4] 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').

values are drawn from the distribution F_G. Formally, $R_G(v) := v \cdot [1 - F_G(v)] \cdot q_G$. 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 $r_G \in \arg\max_p R_G(p)$.*

Definition 2.2 (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 2.3 (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 2.4 (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 $R(\cdot)$.*[5] *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 endpoints of the ironed interval, respectively.*

An ironed revenue curve is depicted in Figure 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 $\beta R(a) + (1-\beta)R(b) = \hat{R}(z)$. One can check that this is precisely the allocation and payment

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

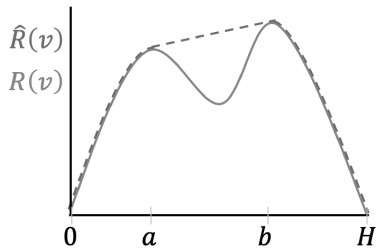


Fig. 1. For some implicit distribution F, the revenue curve $R(v) = v \cdot [1 - F(v)]$ is depicted, as is the ironed revenue curve, or the revenue curve's least concave upper bound.

3. Duality

In this section, we briefly overview the bare minimum duality preliminaries required. Full duality preliminaries are provided in Appendix A.3-A.5.

3.1. Dual Terminology.

In this section, we introduce pictorial representations (Figures 2 and 3) of key aspects of a dual solution and define terminology relevant to the dual.

[5] 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.

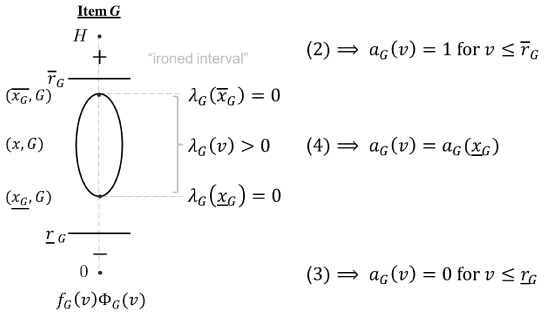


Fig. 2. A pictorial interpretation of virtual values $f_G(v)\Phi_G^{\lambda,\alpha}(v)$ and the dual variable $\lambda_G(v)$, in addition to the concepts of endpoints of the zero region, ironing, an ironed interval, and the allocation in response.

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 2) 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 $\underline{x}_G = \sup\{v \le x \mid \lambda_G(v) = 0\}$ and the top end point $\bar{x}_G = \inf\{v \ge x \mid \lambda_G(v) = 0\}$. 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)) 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.

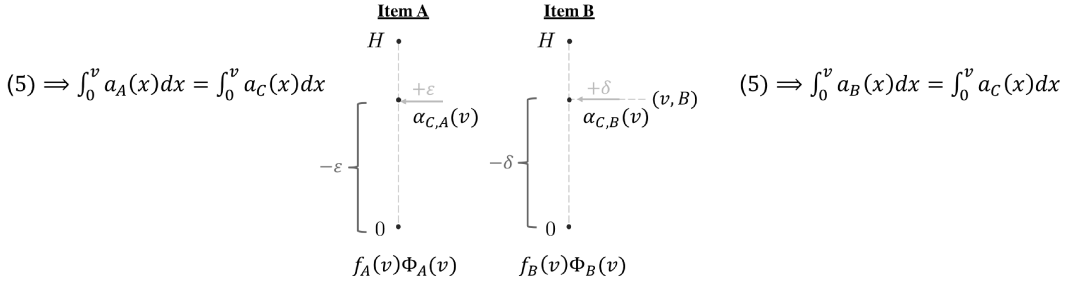


Fig. 3. A pictorial representation of the dual variable α, in addition to the concepts of flow, preferable items, and equally preferable items. Flow is assumed to be coming from item C.

In Figure 3, 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 (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*.[6] Once again, this is a generalization of Myerson's virtual value function to this more general setting.

Figure 2 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 Appendix A.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 2 the horizontal black lines and signs indicate where the virtual values shift from positive sign to zero, $\bar{r}_G$, 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 (2)) and $a_G(v) = 1$ for $v \geq \bar{r}_G$ (condition (3)).

3.2. 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*.

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

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

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

Before continuing, lets 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

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

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.)

3.3. 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:[7]

$$\text{(Primal best response)} \quad \Phi_G^{\lambda,\alpha}(v) > 0 \quad \Rightarrow \quad a_G(v) = 1 \tag{2}$$

$$\Phi_G^{\lambda,\alpha}(v) < 0 \quad \Rightarrow \quad a_G(v) = 0 \tag{3}$$

$$\text{(Dual best response)} \quad \lambda_G(v) > 0 \quad \Rightarrow \quad a'_G(v) = 0 \tag{4}$$

$$\alpha_{G,G'}(v) > 0 \quad \Rightarrow \quad \int_0^v a_G(x)dx - \int_0^v a_{G'}(x)dx = 0 \tag{5}$$

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 paper is using the constraints imposed by complementary slackness in (2-5) to reason about optimal mechanisms and their menu complexity.

4. Menu Complexity

We provide here the key ideas behind the construction that forms our lower bound and the proof of our upper bound. Full details are provided in Appendix E and Appendix F respectively.

4.1. 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 menu complexity lower bound of two, we must additionally:

- Establish that there exists a distribution F for which our dual is feasible. This is not covered in this section, and is deferred to our Master Theorem (Theorem H.1).
- 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 Appendix F.

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 use graphics and recall the following facts established in the previous section:

[7] 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.

(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 $a_G(x) = 0$. (CS2-3)
(2) A "←" 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 $u_A(x) = u_B(x)$. (CS5)
(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)

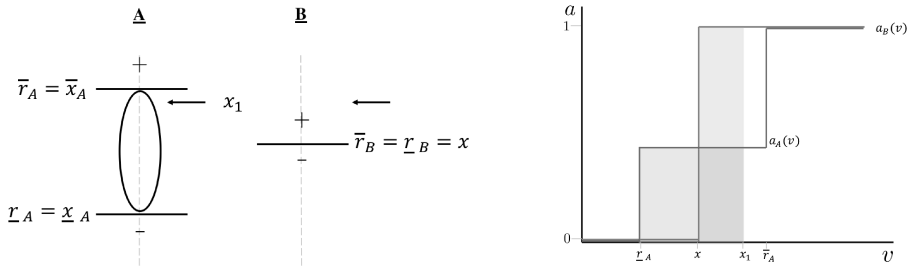


Fig. 4. Left: Our first example that requires randomizing on A, containing an ironed interval $[\underline{r}_A, \overline{r}_A]$ (so $a_A(x_1) = a_A(\underline{r}_A)$) and flow into both A and B at x_1 (so $u_A(x_1) = u_B(x_1)$). Right: Primal best response dictates a price of x for item B, while A's 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 $a_A(\underline{r}_A) \in (0, 1)$.

4.1.1. Step One: the base gadget and a lower bound of $M = 2$. Our base case example is depicted in Figure 4. We note each feature, and how it ties our hands with respect to the allocation rule via complementary slackness.

- 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 (CS2) implies that $a_B(v) = 1$ for $v > x$.
- There is flow into both items A and B at $x_1 > x$. That is, $\alpha_{C,A}(x_1), \alpha_{C,B}(x_1) > 0$. (CS5) 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 $\int_0^{x_1} a_A(w)dw = \int_0^{x_1} a_B(w)dw$, or equivalently $a_A(\underline{r}_A) \cdot (x_1 - \underline{r}_A) = a_B(x) \cdot (x_1 - x)$), 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.

4.1.2. Step Two: two chains and a lower bound of $M = 3$. Our second example (see Figure 5) contains the relevant features from the first example, but extends it to add an additional constraint: we

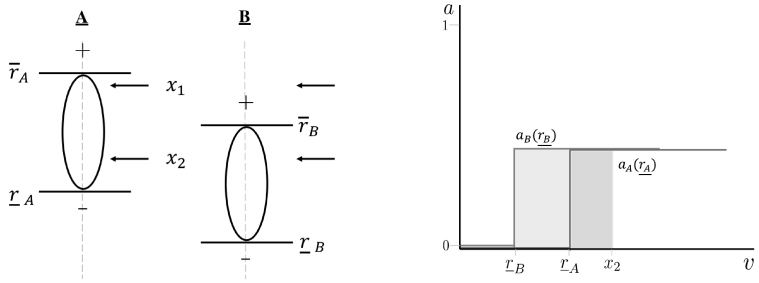


Fig. 5. Left: Our second example, which requires randomization on both A and B.
Right: If $a_A(\underline{r}_A) \le a_B(\underline{r}_B)$, then $u_A(x_2) < u_B(x_2)$ (the blue region is smaller than the red), which violates complementary slackness.

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.

First we observe that the construction requires the allocation rules to be non-zero in the relevant regions, so there is no trivial solution.

- There is flow into both items A and B at $x_1 \in (\overline{r}_B, \overline{r}_A)$: $\alpha_{C,A}(x_1), \alpha_{C,B}(x_1) > 0$. (CS2) 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 $a_B(\underline{r}_B) = a_B(x_2) > 0$.

Now we observe that each of these points' allocations (which we have observed to be non-zero) must be unique.

- 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.

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.

4.1.3. Step Three: four chains and a lower bound of $M = 4$. 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.

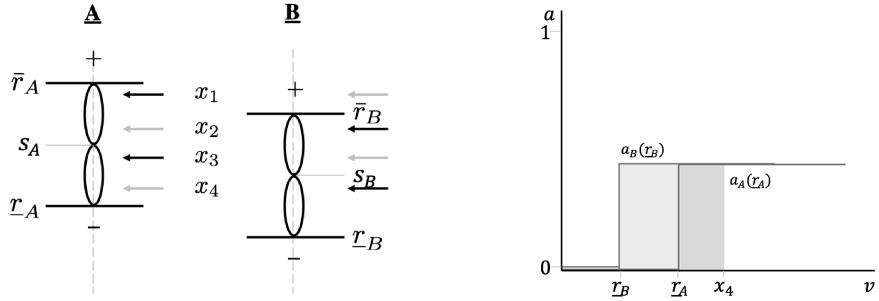


Fig. 6. Left: An optimal dual for our example distributions, which will require at least 4 distinct allocation probabilities.
Right: If $a_A(\underline{r}_A) \le a_B(\underline{r}_B)$, then $u_A(x_2) < u_B(x_2)$ (the blue region is smaller than the red), which violates complementary slackness via Fact 2 at x_4.

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 reiterate the same argument made in the two previous sections, skipping some details.

- Note that $a_B(x_1) > 0$ by Fact 1, and thus $u_B(x_1) > 0$.
- By Fact 2, $u_A(x_1) = u_B(x_1) > 0$. Then $a_A(x_1) > 0$.
- 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 6. 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 $a_A(x_4) > a_B(x_4)$, 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

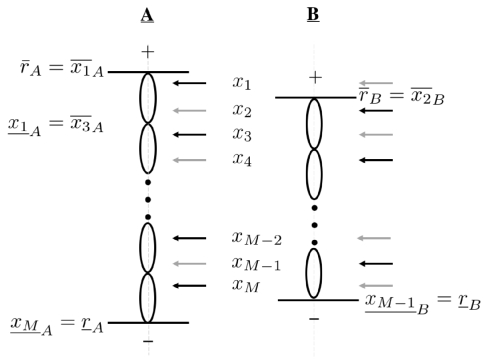


Fig. 7. Our candidate dual instance: a top chain that spans the entire region of zero virtual values for both A and B with no gaps between the ironed intervals that comprise the chain. There is flow into A and B at every point x_i in the chain.

intervals in A and B each with four points of interest. Our full construction below allows the number of ironed intervals to grow with M.

4.1.4. 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 Appendix E, where we formally define this "top chain" structure (Definition E.1) and construct the candidate dual instance, which is depicted in Figure 7. For example, our first example has a top chain of length one, the second of length two, and the third of length four. Theorem 4.2 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 4.1.

Theorem 4.1. *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 F 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 H.1). The formal statement is a bit technical, and can be found in Appendix H.

4.2. For Three Items, Menu Complexity is Finite: Brief Highlight

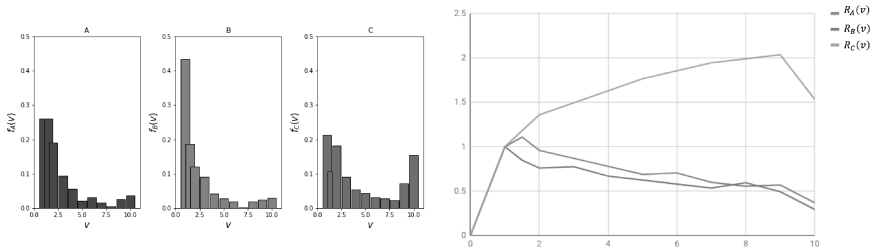
In Appendix F, we discuss our approach for characterizing the optimal mechanism for our 3-item minimal instance. We prove essentially that the interleaving of ironed intervals used in the construction of the previous section is the worst case (in terms of menu complexity). We do this by specifying a 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 4.2. *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 7) 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.

4.3. One Last Example

In this section, we construct an example by applying the Master Theorem (Theorem H.1) to the dual in Figure 6. The customer prior distribution in the example consists of the marginal distributions depicted in Figure 8. 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.



(a) Probability densities

(b) Revenue curves

Fig. 8. The value distributions for items A, B, and C, that do not satisfy DMR.

We can use the revenue curve procedure from Appendix B to determine the optimal pricing for this example. It produces the curves in Figure 8, 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 Section 4.1, for the dual in Figure 6 (which corresponds to this distribution) to satisfy complementary slackness with a mechanism, the mechanism must have a good deal of randomization.

In Section 4.1, 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 4.2 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.

5. Conclusions

We study optimal mechanisms for single-minded bidders, and show that the menu complexity of optimal mechanisms is unbounded but finite for three items. Recall that for three identical items, the menu complexity is 1, for totally-ordered items the menu complexity is at most 7, and for heterogeneous items the menu complexity is uncountable. So our setting fits nicely "in between" totally-ordered and heterogeneous by this measure. By fuzzier measures of complexity, the same is true too: for identical items, the optimal mechanism has a clean closed-form description. For totally-ordered items, the optimal dual has a closed form, and the primal can be recovered by a simple algorithm as a function of this dual. For partially-ordered items, the optimal dual is unlikely to have a closed form, but can be characterized in terms of properties it must satisfy, and the primal can still be recovered algorithmically[8] as a function of this dual. For heterogeneous items, optimal mechanisms are pure chaos. And, like other settings that can be placed fundamentally in between single- and multi- dimensional settings (e.g., FedEx and MUP), we prove that the optimal mechanism is deterministic under DMR in the partially-ordered setting.

We also provide extensions—menu complexity of MUP (Theorem I.1, Appendix I) and of coordinated values (Theorems J.2, J.4, J.5, Appendix J)—proving the usefulness of our techniques beyond our setting.

Many interesting open directions remain. First, general menu complexity upper bounds—for the single-minded setting, the Multi-Unit Pricing setting, and the coordinated valuations setting. The techniques we use in this paper focus on characterizing the optimal dual and recovering the optimal mechanism for the three-item single-minded setting; this approach appears to be far too detailed and focused on characterizations to be extended. We expect new ideas to be needed.

Second, the question of menu-complexity lower bounds for any of these three settings for *approximately*-optimal mechanism are wide-open. Is the separation from FedEx still as large when we only require approximately-optimal revenue?

Both directions of research would further fill out this rich spectrum, which until only recently was but thought to be a dichotomy between single-dimensional and heterogenous.

References

[1] Moshe Babaioff, Yannai A. Gonczarowski, and Noam Nisan. 2017. The Menu-size Complexity of Revenue Approximation. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2017)*. ACM, New York, NY, USA, 869–877. https://doi.org/10.1145/3055399.3055426

[2] Patrick Briest, Shuchi Chawla, Robert Kleinberg, and S Matthew Weinberg. 2015. Pricing lotteries. *Journal of Economic Theory* 156 (2015), 144–174.

[8] Contrast this with [3], which only claims that by solving a linear program, an optimal mechanism for heterogenous settings can be found in time polynomial in the type space.

[3] Yang Cai, Constantinos Daskalakis, and S Matthew Weinberg. 2012. An algorithmic characterization of multi-dimensional mechanisms. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*. 459–478.

[4] Yang Cai, Nikhil R. Devanur, and S. Matthew Weinberg. 2016. A Duality Based Unified Approach to Bayesian Mechanism Design. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing (STOC '16)*. ACM, New York, NY, USA, 926–939. https://doi.org/10.1145/2897518.2897645

[5] Yeon-Koo Che and Ian Gale. 2000. The Optimal Mechanism for Selling to a Budget-Constrained Buyer. *Journal of Economic theory* 92, 2 (2000), 198–233.

[6] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. 2013. Mechanism design via optimal transport. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC '13)*. 269–286. http://doi.acm.org/10.1145/2482540.2482593

[7] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. 2015. Strong duality for a multiple-good monopolist. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*. ACM, 449–450.

[8] Nikhil R. Devanur, Nima Haghpanah, and Christos-Alexandros Psomas. 2017. Optimal Multi-Unit Mechanisms with Private Demands. In *Proceedings of the 2017 ACM Conference on Economics and Computation (EC '17)*. ACM, New York, NY, USA, 41–42. https://doi.org/10.1145/3033274.3085122

[9] Nikhil R. Devanur and S. Matthew Weinberg. 2017. 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)*. ACM, New York, NY, USA, 39–40. https://doi.org/10.1145/3033274.3085132

[10] Amos Fiat, Kira Goldner, Anna R. Karlin, and Elias Koutsoupias. 2016. The FedEx Problem. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC '16)*. ACM, New York, NY, USA, 21–22. https://doi.org/10.1145/2940716.2940752

[11] Yiannis Giannakopoulos and Elias Koutsoupias. 2014. Duality and Optimality of Auctions for Uniform Distributions. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC '14)*. 259–276. http://doi.acm.org/10.1145/2600057.2602883

[12] Yannai A. Gonczarowski. 2018. 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*. 123–131. https://doi.org/10.1145/3188745.3188786

[13] Nima Haghpanah and Jason Hartline. 2015. Reverse mechanism design. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*. ACM, 757–758. Updated version: http://arxiv.org/abs/1404.1341.

[14] Sergiu Hart and Noam Nisan. 2013. The Menu-size Complexity of Auctions. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce (EC '13)*. ACM, New York, NY, USA, 565–566. https://doi.org/10.1145/2482540.2482544

[15] Sergiu Hart and Philip J Reny. 2012. Maximizing revenue with multiple goods: Nonmonotonicity and other observations. Hebrew University of Jerusalem. *Center for Rationality DP-630* (2012).

[16] Jason D Hartline. 2013. Mechanism design and approximation. *Book draft. October* 122 (2013).

[17] Jean-Jacques Laffont, Eric Maskin, and Jean-Charles Rochet. 1987. Optimal nonlinear pricing with two-dimensional characteristics. *Information, Incentives and Economic Mechanisms* (1987), 256–266.

[18] Daniel Lehmann, Liadan Ita O´ Callaghan, and Yoav Shoham. 2002. Truth Revelation in Approximately Efficient Combinatorial Auctions. *J. ACM* 49, 5 (Sept. 2002), 577?602. https://doi.org/10.1145/585265.585266

[19] Alexey Malakhov and Rakesh V Vohra. 2009. An optimal auction for capacity constrained bidders: a network perspective. *Economic Theory* 39, 1 (2009), 113–128.

[20] Alejandro M. Manelli and Daniel R. Vincent. 2007. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory* 137, 1 (2007), 153 – 185. http://dx.doi.org/10.1016/j.jet.2006.12.007

[21] R. Preston McAfee and John McMillan. 1988. Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory* 46, 2 (1988), 335 – 354. http://dx.doi.org/10.1016/0022-0531(88)90135-4

[22] Roger B. Myerson. 1981. Optimal Auction Design. *Mathematics of Operations Research* 6, 1 (1981), 58–73. http://dx.doi.org/10.1287/moor.6.1.58

[23] Raghuvansh R. Saxena, Ariel Schvartzman, and S. Matthew Weinberg. 2018. The Menu-Complexity of One-and-a-Half Dimensional Mechanism Design. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '18)*.

[24] Zihe Wang and Pingzhong Tang. 2014. Optimal Mechanisms with Simple Menus. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation (EC '14)*. ACM, New York, NY, USA, 227–240. https://doi.org/10.1145/2600057.2602863

A. Full Preliminaries

While this paper 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 Devanur and Weinberg [9], Fiat et al. [10]).

A.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. [10], 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.[9] Also observed in Fiat et al. [10] is that Myerson's payment identity holds in this setting as well, and any truthful mechanism must satisfy $p_G(v) = v a_G(v) - \int_0^v a_G(w)dw$ (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:[10]

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

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.

[9] 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.

[10] For the familiar reader, this derivation is routine, so we omit it. The unfamiliar reader can refer to [16, 22] for this derivation.

[10] 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').[11] Below, we will now formulate the Primal LP and its Lagrangian 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.

A.2. The Primal

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

Variables: $a_G(v),\ \forall G \in \mathcal{G},\ v \in [0, H]$

Maximize $$\sum_{G \in \mathcal{G}} \int_0^H f_G(v) a_G(v) \varphi_G(v) dv$$

subject to

$$a'_G(v) \geq 0 \qquad \forall G \in \mathcal{G}\ \forall v \in [0, H] \text{ (dual variables } \lambda_G(v) \geq 0)$$

$$\int_0^v a_G(x)dx - \int_0^v a_{G'}(x)dx \geq 0 \qquad \forall G \in \mathcal{G},\ G' \in N^+(G)\ \forall v \in [0, H] \text{ (dual vars } \alpha_{G,G'}(v) \geq 0)$$

$$a_G(v) \in [0, 1] \qquad \forall G \in \mathcal{G},\ \forall v \in [0, H] \text{ (no dual variables)}$$

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 $[0, 1]$.

A.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

$$\mathcal{L}(a; \lambda, \alpha) :=$$

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

[11] 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').

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:

$$\int_0^H \lambda_G(v)a'_G(v)dv = \lambda_G(v)a_G(v)\ |_0^H - \int_0^H \lambda'_G(v)a_G(v)dv = -\int_0^H \lambda'_G(v)a_G(v)dv$$

As in [FGKK16], 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

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

A.4. More Dual Terminology

Minimal dual terminology is first introduced in subsection 3.1. Here, we add a few additional terms.

Dual best response (condition (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 $u_{G'}(x)=u_{G''}(x)$.

A.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 3.2, 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 [DW17, FGKK16], 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 $v' \le v$, $f_G(v')\Phi_G^{\lambda,\alpha}(v') < f_G(v)\Phi_G^{\lambda,\alpha}(v) < 0$, 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 [DW17]. 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,

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

where $R_C(\cdot)$ is defined as in Definition 2.1, $\hat{R}_C(\cdot)$ 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 [FGKK16].

Theorem A.1 (Proper Ironing [FGKK16]). *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 $f_G(\cdot)\Phi_G^{\lambda,\alpha}(\cdot)$.*

B. 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 [10], 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,[12] 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 Appendix C 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 $r_A := \arg\max_p\{p \cdot [1 - F_A(p)]\}$. 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 9). 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

[12] Recall that a one-dimensional distribution D can stochastically dominate D' yet have a lower Myerson reserve. For example, if D is uniform over the set $\{1, 10\}$, the Myerson reserve is 10. If D is uniform over the set $\{9, 10\}$, the Myerson reserve is 9.

proving it will certainly be technically involved)—without it, there might be a better randomized mechanism.

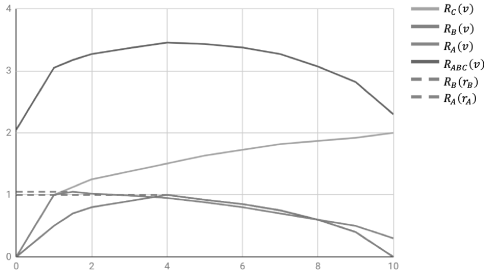


Fig. 9. The construction of R_{ABC} with $R_A, \bar{R}_A, R_B, \bar{R}_B$, and R_C illustrated as well.

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 2.4. 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 10. Figure 11 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 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

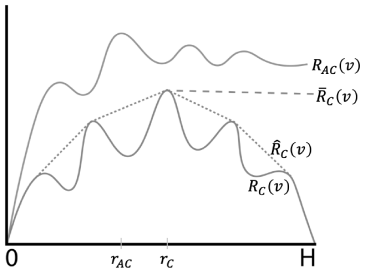


Fig. 10. A worse-to-better item revenue curve for the FedEx setting that determines the optimal mechanism even without DMR.

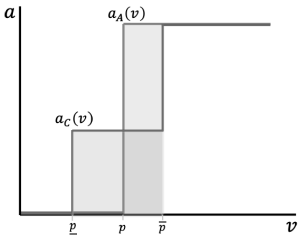


Fig. 11. Utility for items A and C are equal for $v \le \underline{p}$ and $v \ge \underline{p}$, but for $v \in (\underline{p}, \overline{p})$, the randomized option provides more utility.

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 Section 4.1, we explain why it is that the IC constraints can cause the randomization to get so unwieldy, and Section 4.3 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.

C. An Exact Characterization Under the Assumption of DMR

Recall from Subsection 2.2 that when the distributions satisfy DMR, $\lambda_G(v) = 0$ for all (v, G). Our main result in this section is the following:

Theorem C.1. *Consider any partially-ordered preferences for items $\mathcal{G}, \succ$. If the marginal distribution for each item satisfies DMR, the optimal mechanism is deterministic.*

For a deterministic mechanism, we will set a take-it-or-leave-it price p_G for each item G.

C.1. Intuition

It will turn out that the optimal mechanism is analogous to that in FedEx and will set prices as follows:

- For items G that are sink nodes in the DAG, set $p_G = r_G$.
- Starting from the sink nodes and visiting nodes in reverse depth, we will define a least upper bound on each node's price based on the prices set for nodes that dominate it. 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 $N^+(G)$ who are also constrained by this upper bound. Thus, if we follow the sets L_G, we will find 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 paths to the nodes that limit 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 the limiting reserve and never lowers it. We update regularly to ensure that we are always sending flow to the now-limiting reserve, raising it, and thus relaxing 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 flow out at $r_B < v$.

C.2. Formal Pricing Algorithm

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

Dual variable construction:
Base case: For sink nodes A, there is nowhere to send flow. Set $\bar{p}_A = r_A$.
for all nodes A starting from the sink nodes and in increasing reverse depth [13] **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$(A, v, \sigma_A(v))$

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

The key idea is that the price of a node G is limited by the smallest r_A where A is some item better than G (i.e. there is a path from G to A in the DAG). As we send flow along the path to A, we raise r_A and it becomes less limiting. Let S_G be the set of the items that limit G the most, which are precisely the items A such that $r_A = \bar{p}_G$. Since we are in the continuous setting, sending flow is a continuous process. This means that 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 and become less limiting.

Let $L_G \subseteq N^+(G)$ 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 use the variable r to keep track of the updated $\bar{p}_G$. 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 $f_B(r)\Phi_B^{\lambda,\alpha}(r) \le 0$. In every step we decrease the amount of flow to send and the algorithm will terminate when there is no flow left to send. Throughout this process the point r and the set S_G both only increase.

First, we set the flow out of G:

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

Lemma C.2. *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 $f_B(r)\Phi_B^{\lambda,\alpha}(r) = 0$.

[13] i.e. # edges from sink nodes

(3) If $B \in N^+(G) \setminus L_G$, then $r < \bar{r}_B$ and thus $f_B(r)\Phi_B^{\lambda,\alpha}(r) \le 0$.

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

Let ε be 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 $r + \varepsilon < \bar{p}_Z$, 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

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

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

Otherwise, suppose $r + \varepsilon \ge \bar{p}_Z$ and $v \ge \bar{p}_Z$. Then we instead choose $\varepsilon = \bar{p}_Z - r$ and make the same update described above, add Z to L_G and add the item Y that is limiting Z, that is, Y such that $\bar{p}_Z = R_Y$, to S_G. Note that we have sent positive flow, but the flow sent is $< \sigma$. After the update, we will have $r \leftarrow r + \varepsilon = \bar{p}_Z$ and $f_A(r)\Phi_A^{\lambda,\alpha}(r) = 0$ for all $A \in S_G$, including Y. Then again (2) holds, and (3) holds since $r = \bar{p}_Z < \bar{p}_B$ 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 C.3. *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 C.4. *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 $v \le x$ has $f_G(v)\Phi_G^{\lambda,\alpha}(v)$ decrease by α while this remains unchanged for $v > x$, causing no non-monotonicities.

□

We are now ready to prove the main result of this section.

Proof of Theorem C.1. We claim the the following deterministic allocation rule always satisfies complementary slackness with the dual: set $p_G = \min\{r_G, r_A : A \in S_G\}$.

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 C.3 and Lemma C.4.

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 C.2, $f_A(r)\Phi_A^{\lambda,\alpha}(r) = 0$ 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 $v \le \bar{r}_G$ 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 C.2, 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. □

D. An Extension of FedEx: DAGs with Out-Degree At Most 1

In this section, we consider DAGs with out-degree at most 1. That is, partial orders that are tree-like, where each item has at most one item that minimally dominates it. In this case, we see that the FedEx solution applies.

THEOREM D.1. *Consider any partially-ordered preferences for items $\mathcal{G}, \succ$ such that for any G, there exists at most one G'' that minimally dominates G: that is, $G'' \succ G$ and there does not exist any G' where $G'' \succ G' \succ G$. Then a nearly identical construction to the FedEx Problem with a minor modification for partial orderings yields closed-form optimal dual variables and the optimal mechanism.*

We use the notation and methods of [FGKK16]. The proof is almost identical, provided for completeness, and much of the following is duplicated from their paper, with a slight modification to allow for the DAG structure with out-degree at most 1. The key difference is the change in definition of the $\Gamma_{\geq G}$ curves.

We recall the following definitions from their paper:

- Let $\gamma_G(v) := \varphi_G(v) f_G(v)$. Recall that $\varphi_G(\cdot) = v - \frac{1-F_G(v)}{f_G(v)}$.
- Let $\Gamma_G(v) = \int_0^v \gamma_G(x)dx$. As shown in [FGKK16], this function is the negative of the marginal revenue curve for item G. Thus, $\Gamma_G(0) = \Gamma_G(H) = 0$ and $\Gamma_G(v) \leq 0$ for $v \in [0, H]$.
- For any function Γ, define $\hat{\Gamma}(\cdot)$ to be the lower convex envelope [14] 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.
- Let $\hat{\gamma}(\cdot)$ be the derivative of $\hat{\Gamma}(\cdot)$ and let $\gamma(\cdot)$ be the derivative of $\Gamma(\cdot)$.

As shown in [FGKK16], the following facts are immediate from the definition of lower convex envelope:

- $\hat{\Gamma}(v) \leq \Gamma(v) \quad \forall v$.
- $\hat{\Gamma}(v_{\min}) = \Gamma(v_{\min})$ 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).

Now, we redefine the functions $\Gamma_{\geq i}$, which are used to set all of the FedEx dual variables, and can be interpreted as negative combined revenue curves for deadlines i through m. Instead, we redefine them for an item G and all dominated items. In a DAG, we let the set of all source nodes, that is, items that dominate no other items in the partial order, be the set S. Similarly, we call the set of sinks, items that are dominated by no other items, as the set T.

Note that, by assumption, every item has out-degree at most 1. Then the set of items that minimally dominate an item G, $N^+(G)$, is of size 0 or 1. If it is of size 0, then $G \in T$: G is a sink node. That is:

OBSERVATION 1. *For all $G \notin T$, $|N^+(G)| = 1$.*

[14] 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 $\hat{R}(v)$.

For this reason, we define the following notation.

Definition D.1. *For $G \notin T$, let $D(G)$ refer to the single item that minimally dominates G: $G' \in N^+(G)$.*

Now, for any source node $G \in S$, define $\Gamma_{\geq G} = \Gamma_G$. We will define the curves $\Gamma_{\geq G}$ for $G \in \mathcal{G} \setminus G$ inductively. Let define $r_{\geq G} := \max \arg\min \Gamma_{\geq G}(v)$. Then define

$$\bar{\Gamma}_{\geq G}(v) := \begin{cases} \hat{\Gamma}_{\geq G}(v) & v < r_{\geq G} \\ \Gamma_{\geq G}(r_{\geq G}) & v \geq r_{\geq G}. \end{cases}$$

Now, for all G that are not source nodes, we can inductively define

$$\Gamma_{\geq G}(v) := \Gamma_G(v) + \sum_{G' \in N^-(G)} \bar{\Gamma}_{\geq G'}(v).$$

Note then that

$$\bar{\gamma}_{\geq G} = \begin{cases} \hat{\gamma}_{\geq G} & v \leq r_{\geq G} \\ 0 & v > r_{\geq G} \end{cases} \qquad \text{and} \qquad \gamma_{\geq G} = \gamma_G + \sum_{G' \in N^-(G)} \bar{\gamma}_{\geq G'}(v).$$

D.1. Primal, Dual, and Complementary Slackness

We use the following primal and dual formulations with the noted complementary slackness conditions. They are virtually identical to FedEx, modified for the DAG, and much of it is copied from [FGKK16].

The Primal

Variables: $a_G(v)$, for all $G \in \mathcal{G}$, and all $v \in [0, H]$.

$$\text{Maximize} \sum_{\mathcal{G}} \int_0^H a_G(v)\gamma_G(v)dv$$

Subject to

$$\begin{aligned} \int_0^v a_G(x)dx - \int_0^v a_{G'}(x)dx \leq 0 \qquad & \forall G \in \mathcal{G} \setminus S, G' \in N^-(G) \quad \forall v \in [0,H] \qquad \text{(dual variables } \alpha_{G',G}(v)\text{)} \\ a_G(v) \leq 1 \qquad & \forall G \in \mathcal{G} \quad \forall v \in [0,H] \qquad \text{(dual variables } b_G(v)\text{)} \\ -a'_G(v) \leq 0 \qquad & \forall G \in \mathcal{G} \quad \forall v \in [0,H] \qquad \text{(dual variables } \lambda_G(v)\text{)} \\ a_G(v) \geq 0 \qquad & \forall G \in \mathcal{G} \quad \forall v \in [0,H]. \end{aligned}$$

Note that $a'_G(v)$ denotes $\frac{d}{dv} a_G(v)$.

The Dual

Variables: $b_G(v), \lambda_G(v)$, for all $G \in \mathcal{G}$, and all $v \in [0,H]$, $\alpha_{G',G}(x)$ for $G \in \mathcal{G} \setminus S, G' \in N^-(G)$ and all $x \in [0,H]$.

$$\text{Minimize} \int_0^H \sum_{G \in \mathcal{G}} b_G(v) dv$$

Subject to

$$b_G(v) + \lambda'_G(v) + \sum_{G' \in N^-(G)} \int_v^H \alpha_{G',G}(x) dx \geq \gamma_G(v) \qquad \forall v \in [0,H], G \in S \text{ (primal var } a_G(v))$$

$$b_G(v) + \lambda'_G(v) + \sum_{G' \in N^-(G)} \int_v^H \alpha_{G',G}(x) dx \qquad \forall v \in [0,H], i \in \mathcal{G} \setminus S, T$$

$$- \int_v^H \alpha_{G,D(G)}(x) dx \geq \gamma_G(v) \qquad \text{(primal var } a_G(v))$$

$$b_G(v) + \lambda'_G(v) - \int_v^H \alpha_{G,D(G)}(x) dx \geq \gamma_G(v) \qquad \forall v \in [0,H], G \in T \text{ (primal var } a_G(v))$$

$$\lambda_G(H) = 0 \qquad \forall G \in \mathcal{G}$$

$$\alpha_{G',G}(v) \geq 0 \qquad \forall v \in [0,H], G \in \mathcal{G} \setminus S, G' \in N^-(G)$$

$$b_G(v), \lambda_G(v) \geq 0 \qquad \forall G \in \mathcal{G} \forall v \in [0,H].$$

Note that $\lambda'_G(v)$ denotes $\frac{d}{dv}\lambda_G(v)$.

D.2. Conditions for strong duality

As long as there are feasible primal and dual solutions satisfying the following conditions, strong duality holds. Theorem 3 from [FGKK16] proves that these conditions are sufficient.

$$a_G(v) > 0 \quad \Rightarrow \quad \lambda_G(v) \text{ continuous at } v \quad G \in \mathcal{G} \tag{6}$$

$$a_G(v) < 1 \quad \Rightarrow \quad b_G(v) = 0 \qquad G \in \mathcal{G} \tag{7}$$

$$a'_G(v) > 0 \quad \Rightarrow \quad \lambda_G(v) = 0 \qquad G \in \mathcal{G} \tag{8}$$

$$\int_0^v a_G(x) dx < \int_0^v a_{G'}(x) dx \quad \Rightarrow \quad \alpha_{G',G}(v) = 0 \; G \in \mathcal{G} \setminus S, G' \in N^-(G) \tag{9}$$

$$b_G(v) + \lambda'_G(v) + \sum_{G' \in N^-(G)} \int_v^H \alpha_{G',G}(x) dx$$

$$- \int_v^H \alpha_{G,D(G)}(x) dx > \gamma_G(v) \quad \Rightarrow \quad a_G(v) = 0 \qquad G \in \mathcal{G} \setminus S, T \tag{10}$$

$$b_G(v) + \lambda'_G(v) + \sum_{G' \in N^-(G)} \int_v^H \alpha_{G',G}(x) dx > \gamma_G(v) \quad \Rightarrow \quad a_G(v) = 0 \qquad G \in T \tag{11}$$

$$b_G(v) + \lambda'_G(v) - \int_v^H \alpha_{G,D(G)}(x) dx > \gamma_G(v) \quad \Rightarrow \quad a_G(v) = 0 \qquad G \in S \tag{12}$$

We allow $a'_G(v) \in \mathbb{R} \cup \{+\infty\}$. It may have (countably many) discontinuities. However, the proof of optimality in [FGKK16] handles this.

D.3. Optimal Primal Variables

We determine the allocation rules inductively, from sink nodes all the way to source nodes. First, for sink nodes $G \in T$, set

$$a_G(v) = \begin{cases} 0 & v < r_{\geq G} \\ 1 & v \geq r_{\geq G}. \end{cases}$$

Suppose that $a_{G'}$ has been defined for $G' = D(G)$, 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,

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

Thus, we can write

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

where

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

Next we define $a_G(v)$.

Definition D.2. *Let j^* be the largest j such that $v_j \leq r_{\geq G}$. For any $j \leq j^*$, consider two cases:*

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

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

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

  Let $0 < \delta < 1$ such that

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

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

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

  Define

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

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

$$a_G(v) = \begin{cases} \sum_{j=1}^{j^*} (\beta_j - \beta_{j-1}) a_{G,j}(v) & \text{if } v < r_{\geq G}, \\ 1 & v \geq r_{\geq G}. \end{cases} \tag{13}$$

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

D.4. Closed-Form Dual Variables and Proof of Optimality

The following dual variables and proofs are again almost verbatim from [FGKK16] with very small modifications for the DAG structure.

$$\lambda_G(v) = \Gamma_{\geq G}(v) - \hat{\Gamma}_{\geq G}(v) \tag{14}$$

$$b_G(v) = \begin{cases} 0 & v < r_{\geq G} \\ \hat{\gamma}_{\geq G}(v) & v \geq r_{\geq G} \end{cases} \tag{15}$$

$$\alpha_{G',G}(v) = \begin{cases} \hat{\gamma}'_{\geq G'}(v) & v < r_{\geq G'} \\ 0 & v \geq r_{\geq G'} \end{cases} \qquad \text{for } G \in \mathcal{G} \setminus S, G' \in N^-(G) \tag{16}$$

Taking the derivative of (14), and using the definition of Γ, we obtain:

$$\gamma_G(v) - \lambda'_G(v) = \hat{\gamma}_{\geq G}(v) - \sum_{G' \in N^-(G)} \bar{\gamma}_{\geq G'}(v) \qquad \text{for } G \in \mathcal{G} \setminus S \tag{17}$$

$$\gamma_G(v) - \lambda'_G(v) = \hat{\gamma}_G(v) \qquad \text{for } G \in S \tag{18}$$

Also, using (16) and the fact that $\hat{\gamma}_{\geq i+1}(r_{\geq i+1}) = 0$, we get:

$$A_{G',G}(v) := \int_v^H \alpha_{G',G}(x)\, dx = \begin{cases} -\hat{\gamma}_{\geq G'}(v) & v < r_{\geq G'} \\ 0 & v \geq r_{\geq G'} \end{cases} = -\bar{\gamma}_{\geq G'}(v) \tag{19}$$

Claim. *Conditions (10)–(12) and dual feasibility: For all G and v, $a_G(v) > 0 \implies$ the corresponding dual constraint is tight, and the dual constraints are always feasible.*

Proof. Rearrange the dual constraint $b_G(v) + \sum_{G' \in N^-(G)} A_{G',G}(v) - A_{G,D(G)}(v) + \lambda'_G(v) \geq \gamma_G(v)$ to

$$b_G(v) - A_{G,D(G)}(v) \geq \gamma_G(v) - \lambda'_G(v) - \sum_{G' \in N^-(G)} A_{G',G}(v).$$

Fact 1: For $G \notin S$, $\gamma_G(v) - \lambda'_G(v) - \sum_{G' \in N^-(G)} A_{G',G}(v) = \hat{\gamma}_{\geq G}(v)$ for all v. To see this use (17) and (19):

$$\gamma_G(v) - \lambda'_G(v) = \hat{\gamma}_{\geq G}(v) - \sum_{G' \in N^-(G)} \bar{\gamma}_{\geq G'}(v) \qquad A_{G',G}(v) = -\bar{\gamma}_{\geq G'}$$

Fact 2: For $G \notin T$, $b_G(v) - A_{G,D(G)}(v) = \hat{\gamma}_{\geq G}(v)$ for all v.

$$b_G(v) = \begin{cases} 0 & v < r_{\geq G} \\ \hat{\gamma}_{\geq G}(v) & v \geq r_{\geq G} \end{cases} \qquad -A_{G,D(G)}(v) = \begin{cases} \hat{\gamma}_{\geq G}(v) & v < r_{\geq G} \\ 0 & v \geq r_{\geq G} \end{cases}$$

Hence for $G \notin T$, $b_G(v) - A_{G,D(G)}(v) \geq \gamma_G(v) - \lambda'_G(v) - \sum_{G' \in N^-(G)} A_{G',G}(v)$ for all v.

For $G \in S$, since $\gamma_{\geq G} = \gamma_G$, and $\gamma_G(v) - \lambda'_G(v) = \hat{\gamma}_G(v)$. Combining this with Fact 2 above, we get that $b_G(v) - A_{G,D(G)}(v) + \lambda'_G(v) = \gamma_G(v)$ for all v.

Finally, for $G \in T$, using Fact 1, for $v < r_{\geq G}$, we get

$$b_G(v) = 0 \geq \hat{\gamma}_{\geq G}(v) = \gamma_G(v) - \lambda'_G(v) - \sum_{G' \in N^-(G)} A_{G',G}(v)$$

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

$$b_G(v) = \gamma_{\geq G}(v) = \gamma_G(v) - \lambda'_G(v) - \sum_{G' \in N^-(G)} A_{G',G}(v),$$

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

The proofs of conditions Appendix D.2–Appendix D.2 are identical as in [FGKK16] using our slightly modified dual variables:

Claim. *Condition (7): For all G and v, $a_G(v) < 1 \implies b_G(v) = 0$.*

Proof. If $a_G(v) < 1$, then $v < r_{\geq G}$, so by construction, $b_G(v) = 0$. □

Claim. *Condition (8): For all G and v, $a'_G(v) > 0 \implies \lambda_G(v) = 0$.*

Proof. From Subsection 4.2 of [FGKK16], $a'_G(v) > 0$ only for unironed values of v, at which $\lambda_G(v) = 0$. □

Claim. *Condition (9): For all $G \in \mathcal{G} \setminus S, G' \in N^-(G)$ and v, $\int_0^v a_G(x)dx < \int_0^v a_{G'}(x)dx \implies \alpha_{G',G}(v) = 0$.*

Proof. As discussed at the end of the proof of Lemma 1 of [FGKK16], $\int_0^v a_G(x)dx = \int_0^v a_{G'}(x)dx$ unless $\Gamma_{\geq G'}$ is ironed at v, or $v \geq r_{\geq G}$. In both of these cases $\alpha_{G',G}(v) = 0$ (by our fourth fact about the lower convex envelope and equation Equation (16), respectively). □

The above claims prove that this dual solution satisfies feasibility and all complementary slackness and continuity conditions from Section D.2 hold.

E. Details from Section 4.1: Construction of a Candidate Dual Instance

We extend the above examples from Section 4.1 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 12. 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).

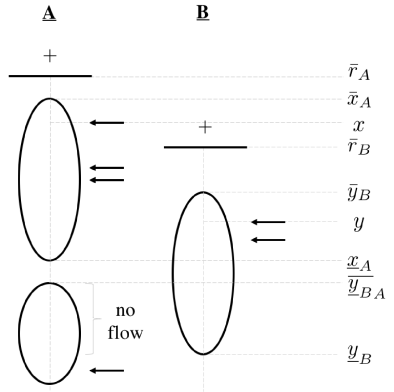


Fig. 12. This is an example of a chain that consists of the points $\{(x,A),(y,B)\}$. It is a top chain as $x > \bar{r}_B$. Note that (y,B) is preceded by (x,A) as there is flow into B at y and $y > \underline{x}_A$, and there is no flow into B for any $v \in (y,\bar{y}_B]$. The chain terminates at (y,B) since there is no flow into A for any $v \in [\underline{y}_B, \overline{\underline{y}_{BA}}]$.

Definition E.1 (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 4.1. First, we provide a construction of our candidate dual, which is depicted in Figure 7. The instance uses definition E.1 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 1. *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 2. *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 3. *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 E.1. *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 1 and Remark 2. 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 3), 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 4.1.

Proof of Theorem 4.1. 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 3), we must have $a_B(x_M) > a_A(x_M)$. However, with this fact, Lemma E.1 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 4.2, 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 E.2. *This idea gives a lower bound for Multi-Unit Pricing as well.*

We expand on this on Appendix I.

F. 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 F.1 (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 F.2. *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 13.*

Definition F.3. *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 14.*

Proposition F.1. *The best dual has no double swaps or upper swaps.*

The full proof of Proposition F.1 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 F.1) and proves that (2) or (3) respectively must not have held for the original dual.

Theorem 4.2. *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 4.2. (No bad structures exist in best duals.) First, we try to satisfy the necessary complementary slackness system of equations as in Appendix E, 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 F.1.

(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 17 in Appendix F.) 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 Appendix E, 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 Appendix E 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 [9, 10]. 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 F.2 ([9]). *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 $G' > G$.*
- *(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 F.1. By Definition F.1, 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 F.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 13.

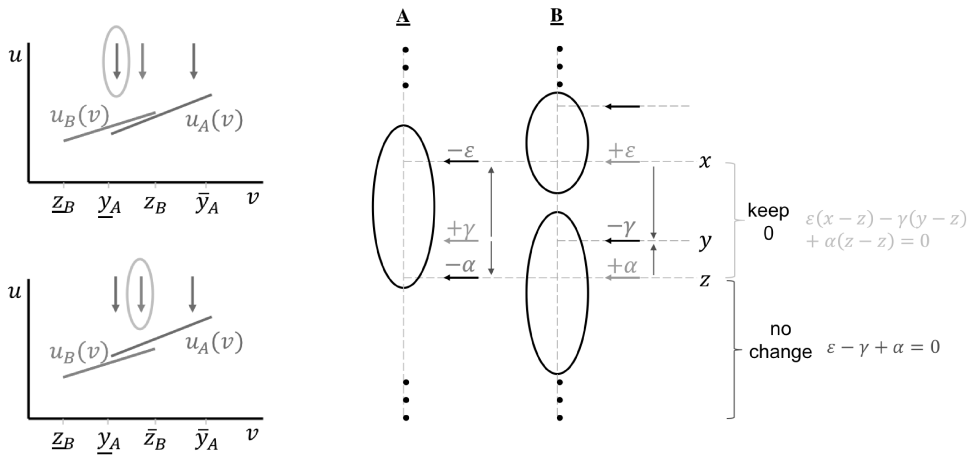


Fig. 13. Left: Complementary slackness forces linear utility in ironed intervals. For any choice of linear utility functions, we cannot satisfy the preferability constraints imposed by the double swap for item A, then B, then A in this region. The violated constraint corresponds to the circled arrow. Right: The operation used in the proof of Lemma F.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 13) 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 $v \le z$, $\hat{\Phi}_A^{\lambda,\alpha}(v) = \Phi_A^{\lambda,\alpha}(v)$ 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.

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

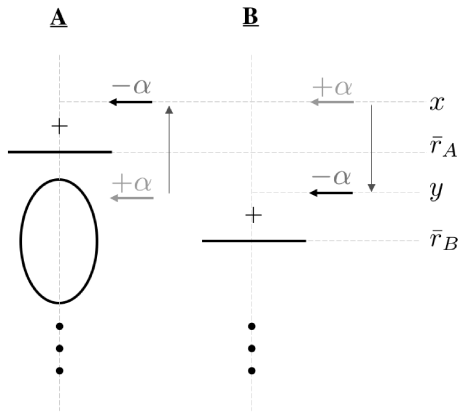


Fig. 14. The operation used in the proof of Lemma F.4, using a upper swap to maintain virtual welfare and create lower positives.

$$= \int_z^x f_A(v)\Phi_A^{\lambda,\alpha}(v)dv + \varepsilon(x-z) - \gamma(y-z)$$
$$= \int_z^x f_A(v)\Phi_A^{\lambda,\alpha}(v)dv$$

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 F.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 14) 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 $\alpha(x-y)$. 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 F.3 and Lemma F.4 comprise the proof of Proposition F.1. □

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 4.2. 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.

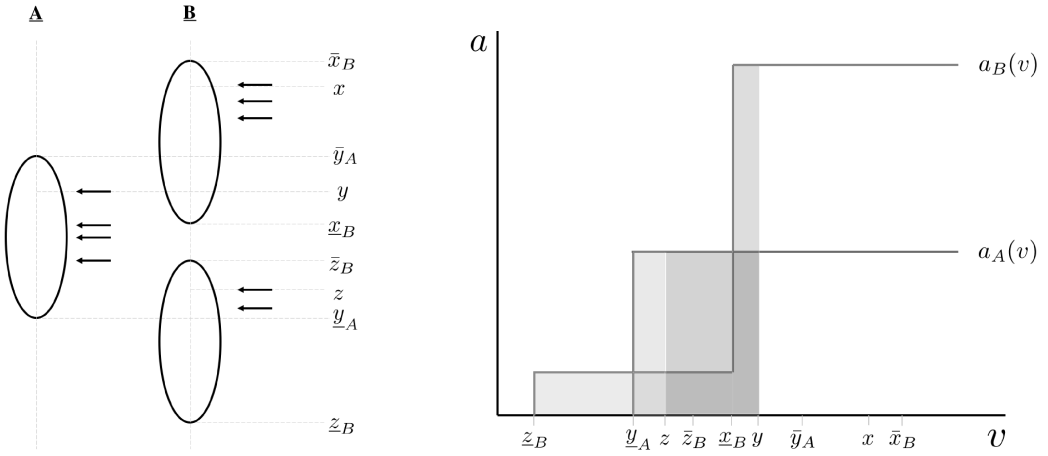


Fig. 15. Left: A candidate dual (with no double or upper swaps); part of a chain. Right: An allocation that satisfies complementary slackness up to value y, satisfying equal preferability at z and y and preferability at all points with flow in.

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

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

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 $a_A(\underline{x}_A) =$

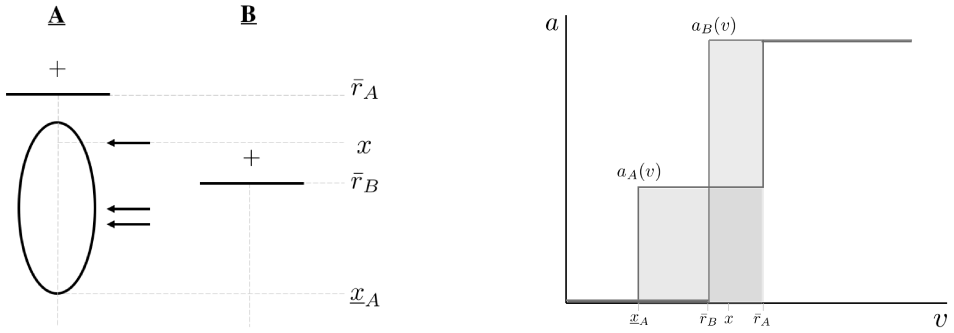


Fig. 16. Left: The base case of a candidate dual with an empty chain. Right: An allocation that satisfies complementary slackness.

$a_B(\bar{r}_B) = 1$. We construct an allocation rule $\hat{a}$ for the top chain of size $n+1$ as follows; this is depicted in Figure 17. Let $\lambda = \frac{x-\bar{r}_B}{x-y-a_B(\underline{y}_B)(\bar{r}_B-y)} < 1$. Then let

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

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. *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.

□

□

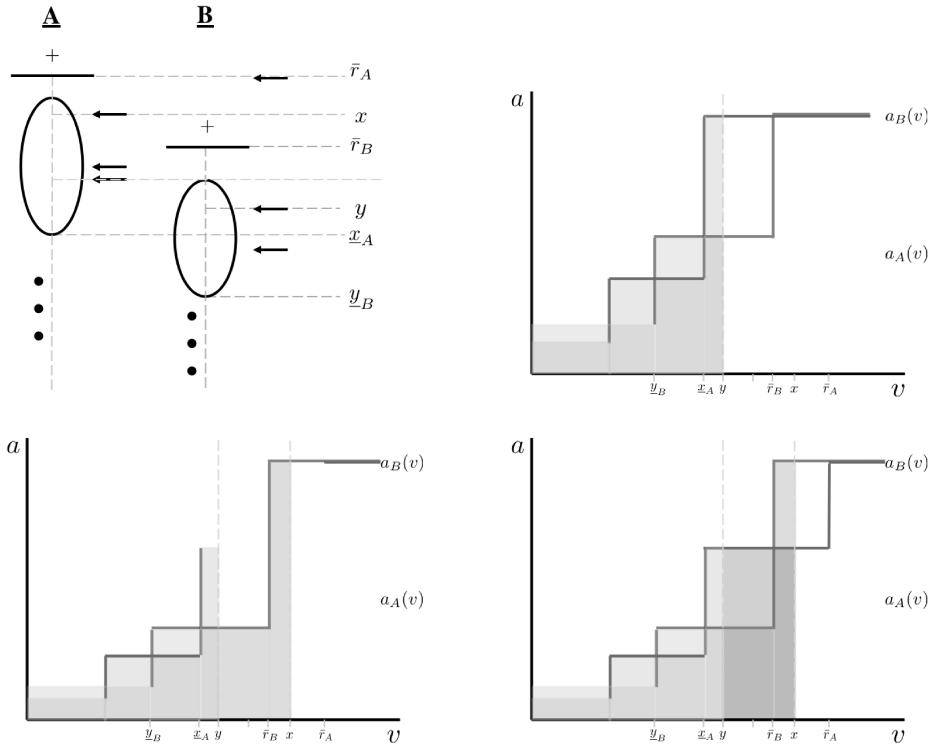


Fig. 17. Top Left: A top chain from a candidate dual. We use the inductive hypothesis on the chain of one size smaller (below the green line). Top Right: The allocation rule from the inductive hypothesis that satisfies all CS constraints on the smaller chain (below the green line). Bottom Left: The scaled allocation rule, requiring preferability of A between the green lines. Bottom Right: The allocation rule that satisfies these preferability constraints.

G. Equivalence with Single-Minded Valuations

In the introduction, we note the following observation.

Observation 2. *The partially-ordered setting is* equivalent *to the single-minded setting.*

First, we define the single-minded setting.

Definition G.1. *In a* single-minded setting*, a seller determines how to sell any bundle of k items. A buyer has a (value, bundle) pair (v, B) where $B \in 2^{[k]}$ is any subset of items. The pair (v, B) is drawn from a joint probability distribution over $[0, H] \times 2^{[k]}$ 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, $2^{[k]}$. 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 $B \in 2^{[k]}$ which do not have an interest that maps onto it, assign measure 0 to the event of drawing (v, B) from F. Otherwise, $f_{B(G)}(v) = f_G(v)$.

H. 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 H.1 (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 $\underline{r}_G \le \underline{x}_{i,G} \le \overline{x}_{i,G} \le \bar{r}_G$, 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 H.1. On our way to prove this theorem, we generalize a result of [23], 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 [23] from discrete distributions to continuous distributions.

Lemma H.2 (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 H.3 (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 H.2. 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 2.1). 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 4. *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 H.1. If the constraint over flows wasn't there, the problem would be a direct application of Theorem H.3. 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 $y \le y_i$ will increase by ε and all points of item G' with $y \le y_i$ 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 $\varepsilon = \frac{1}{2H}$) from the "suggested" (by Theorem H.3) 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 H.4 (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 $1 \le i \le n$ 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 $x_{-i} \le \underline{x}_{j,i} \le \overline{x}_{j,i} \le x_{+i}$. 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 $\alpha_{G_i, G_{i+1}}(y_{j,i,i+1}) \ge 0$ (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 $\alpha_{G_i, G_{i-1}}(y_{j,i,i-1}) \ge 0$ (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 H.1 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 H.1 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. □

I. 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 [DHP17], 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 E.1) 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 $(x_1, A), (x_2, B), \ldots, (x_M, A)$ 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:

- $\underline{r}_A = \underline{x_M}_A/2$, $\bar{r}_A = \overline{x_1}_A/2$, and for i such that (x_i, A) and (x_{i+2}, A) are in the chain, $\underline{x_i/2}_A = \underline{x_{i+2}/2}_A$.
- $\underline{r}_B = \underline{x_M}_B$, $\bar{r}_B = \overline{x_2}_B$, 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 $v \geq \bar{r}_G$; $f_G(v)\Phi_G^{\lambda,\alpha}(v) = 0$ for $v \in [\underline{r}_G, \bar{r}_G]$; $f_G(v)\Phi_G^{\lambda,\alpha}(v) < 0$ 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 $\bar{r}_C = 2\bar{r}_A$.

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 $\alpha_{C,B}(x \to x) > 0$.

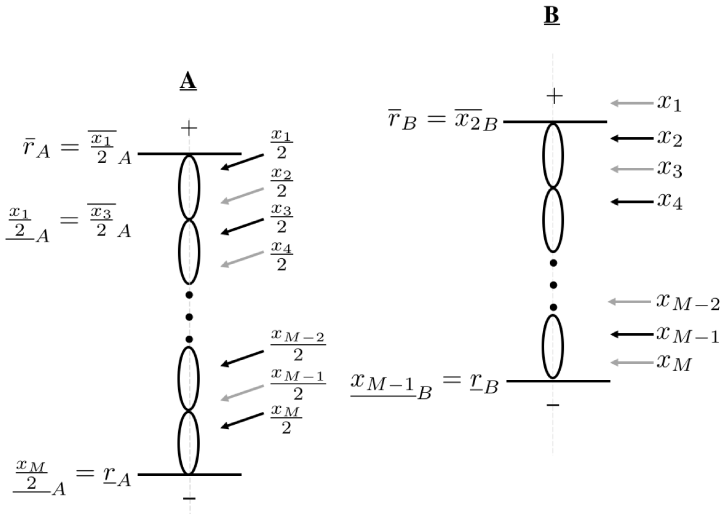


Fig. 18. The analogue of the partially ordered candidate dual, adjusted for the Multi-Unit Pricing problem.

Theorem I.1. *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 4.1. 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 $\geq M$. □

J. Coordinated Valuations

In this section and the following Appendix K, we examine the same minimal partial ordering with $\mathcal{G} = \{A, B, C\}$ where $A \succ C, B \succ C, A \not\succ B, B \not\succ C$. However, a type (v, C) now has a function $g_A(v)$ and $g_B(v)$ describe his valuations for A and B respectively, and $g_C(v) = v$. That is, if a buyer with type (v, C) gets item G, his utility is $g_G(v)$ times the probability that he is served minus his payment.

The main result is that even for g_A that is piecewise linear with only two segments and for $g_B(v) = v$, the randomization required in the optimal mechanism jumps from unbounded but finite, as it was in the partially-ordered setting, to *countably infinite*. This further fills out the spectrum, placing this setting between partially-ordered items and two additive items.

However, if g_A and g_B are not piecewise linear, the menu complexity once again jumps up, becoming uncountably infinite, and matching the menu complexity for two additive items.

J.1. Preliminaries

Consider selling 3 items A, B, and C to one bidder. Define the set Items $= \{A, B, C\}$. We use $\bar{G}$ to refer to a general item $\in$ Items. When we make claims about $\bar{G}$, we mean that the claim holds for each of the three items in Items. When referring only to items A, B, we use the symbol G. Thus, a claim that holds for G holds for both A and B.

In the setting we consider, the bidder has a type in the set Types $= \{A, B, C\} \times [0, H]$, where $H \in \mathbb{R}$ is some constant. The type $(\bar{G}, v)$ of the bidder is drawn from a distribution $\mathfrak{f}$ supported on Types. Denote by $q_{\bar{G}} = \int_0^H \mathfrak{f}(\bar{G}, t)dt$. Also, define $f_{\bar{G}}(v) = \frac{1}{q_{\bar{G}}} \int_0^v \mathfrak{f}(\bar{G}, t)dt$ as the distribution $\mathfrak{f}$ conditioned on $q = \bar{G}$. Almost exclusively, we refer to $\mathfrak{f}$ as $(q, f_{\bar{G}})$. We also omit the subscript when it is clear from context.

We now define the bidder's value function $\mathfrak{v}$: Types $\times$ Items $\to [0, H]$. This is defined as

$$\mathfrak{v}(C, v, \bar{G}) = g_{\bar{G}}(v) \qquad \text{and} \qquad \mathfrak{v}(G, v, \bar{G}) = v \cdot \mathbb{1}_{G=\bar{G}},$$

where $g_G : [0, H] \to [0, H]$ is an increasing, invertible function that is Lipschitz with parameter L and $g_C(v) = v$ for all $v \in [0, H]$. Intuitively, this means that if the bidder has type (C, v), then their value for item $\bar{G}$ is $g_{\bar{G}}(v)$. A bidder with type (G, v) has the value 0 for any item $\bar{G} \neq G$ and value v for the item G.

Mechanisms. A mechanism is defined by two functions a, p, where a : Types $\to [0, 1]$ and p : Types $\to [0, H]$. The function a is called the *allocation rule* and the function p is called the *payment rule*. We will use $a_{\bar{G}}(v)$ to denote $a(\bar{G}, v)$ and $p_{\bar{G}}(v)$ to denote $p(\bar{G}, v)$. A mechanism is said to be *incentive compatible* if, for all $\mathrm{t} = (\bar{G}, v), \mathrm{t}' = (\bar{G}', v') \in$ Types

$$\mathfrak{v}(\mathrm{t}, \bar{G})a(\mathrm{t}) - p(\mathrm{t}) \geq \mathfrak{v}(\mathrm{t}, \bar{G}')a(\mathrm{t}') - p(\mathrm{t}') \tag{20}$$

Instance. An instance $\mathcal{I}$ for the coordinated items setting is defined by a tuple $(q, f_{\bar{G}}, g_{\bar{G}})$. We will usually omit the subscript and simply write (q, f, g). Our goal is find, for a given instance $\mathcal{I}$, the incentive compatible mechanism that maximizes the *revenue* $\mathbb{E}_{\mathrm{t}\sim(q,f)}[p(\mathrm{t})]$.

J.1.1. A Linear Programming Formulation For an instance $\mathcal{I}$ for the coordinated items , finding the incentive compatible mechanism that maximizes $\mathbb{E}_{\mathrm{t}\sim(q,f)}[p(\mathrm{t})]$ turns out to be equivalent to the linear program in Equation 21. Here and throughout, for a distribution f supported on $[0,H]$, we use $\varphi(v) = v - \frac{1-F(v)}{f(v)}$ to denote the Myerson's virtual value function, where $F(v) = \int_0^v f(t)dt$ is the cumulative distribution function for the distribution defined by f.

$$\text{maximize} \qquad \mathcal{P}(a) = \sum_{\bar{G}} \int_0^H q_{\bar{G}} f_{\bar{G}}(t) a_{\bar{G}}(t) \varphi_{\bar{G}}(t) dt \tag{21a}$$

subject to

$$\int_0^{g_G(v)} a_G(t)dt - \int_0^v a_C(t)dt \le 0 \qquad ,\forall v \tag{21b}$$

$$-a'_{\bar{G}}(v) \le 0 \qquad ,\forall v \tag{21c}$$

$$0 \le a_{\bar{G}}(v) \le 1 \qquad ,\forall v \tag{21d}$$

$$a_{\bar{G}}(0) = 0 \tag{21e}$$

Even though the argument that Equation 21 is equivalent to finding the incentive compatible mechanism that maximizes $\mathbb{E}_{\mathrm{t}\sim(q,f)}[p(\mathrm{t})]$ is standard, we summarize it here. Observe that the non-trivial incentive compatibility constraints in Equation 20 can be classified into two types:

- $\mathrm{t} = (\bar{G}, v)$ **and** $\mathrm{t}' = (\bar{G}, v')$**:** In this case, Equation 20 reduces to $v a_{\bar{G}}(v) - p_{\bar{G}}(v) \ge v a_{\bar{G}}(v') - p_{\bar{G}}(v')$. This is attainable for all v, v' if and only if the allocation rule $a_{\bar{G}}$ is monotone increasing (Equation 21c) and $p_{\bar{G}}(v) = v a_{\bar{G}}(v) - \int_0^v a_{\bar{G}}(t)dt$.
- $\mathrm{t} = (C, v)$ **and** $\mathrm{t}' = (G, v')$**:** In this case, Equation 20 reduces to $v a_C(v) - p_C(v) \ge g_G(v) a_G(v') - p_G(v')$. Due to the constraints in the previous paragraph, it is sufficient to have $v a_C(v) - p_C(v) \ge g_G(v) a_G(g_G(v)) - p_G(g_G(v))$. This is equivalent to the constraint in Equation 21b.

Finally, it can be verified that if $p_{\bar{G}}(v) = v a_{\bar{G}}(v) - \int_0^v a_{\bar{G}}(t)dt$, then $\mathbb{E}_{\mathrm{t}\sim(q,f)}[p(\mathrm{t})] = \mathcal{P}(a)$.

We will use $\mathfrak{p} = (a_{\bar{G}})$ to denote a general solution to Equation 21 and $\mathcal{P}(\mathcal{I})$ to denote the optimal value for the instance $\mathcal{I}$.

J.1.2. A Lagrangian Dual Formulation For any instance $\mathcal{I}$ of coordinated items , the revenue maximization problem is defined in Equation 21. Let $X \subset [0,H]$ be a discrete set of points in $[0,H]$. Define the X-dual of Equation 21 as:

$$\text{minimize } \mathscr{D}_X(\lambda, \gamma, \Gamma) = \sum_{\bar{G}} \int_0^H q_{\bar{G}} f_{\bar{G}}(t) \max\left(0, \Phi_{\bar{G}}^{\lambda,\gamma,\Gamma,X}(t)\right) dt \tag{22a}$$

subject to

$$\lambda_{\bar{G}}(H) = 0 \tag{22b}$$

$$\lambda_{\bar{G}}(v), \gamma_G(v), \Gamma_G(x) \ge 0 \qquad ,\forall x \in X, v \in [0,H] \tag{22c}$$

where

$$\Phi_{\bar{G}}^{\lambda,\gamma,\Gamma,X}(v) = \varphi_{\bar{G}}(v) + \frac{1}{q_{\bar{G}} f_{\bar{G}}(v)} \left[-\lambda'_{\bar{G}}(v) - \sum_{X \ni x > g_{\bar{G}}^{-1}(v)} \Gamma_{\bar{G}}(x) - \int_{g_{\bar{G}}^{-1}(v)}^{H} \gamma_{\bar{G}}(s) ds \right]$$

Equation 22 is obtained from Equation 21 by Lagrangifying the constraints Equation 21c using the variables $\lambda_{\bar{G}}(v)$, the constraints Equation 21b using the variables $\gamma_G(v)$, and the constraints Equation 21b, for $v \in X$ using the variables $\Gamma_G(v)$. Here and throughout this paper, we define $\gamma_C(v) = -\gamma_A(v) - \gamma_B(v)$ and $\Gamma_C(v) = -\Gamma_A(v) - \Gamma_B(v)$ for notational ease. We note that getting Equation 22 from Equation 21 requires integrating a certain term by parts. *Throughout this work, we assume that our functions are well-behaved enough to allow such standard operations.*

Note that the idea of using a Lagrangian dual is not new to this work. Indeed, a lot of recent advances in similar settings have been made using this technique. We will use $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ to denote a general solution to Equation 22 and $\mathscr{D}_X(\mathcal{I})$ to denote the optimal value for the instance $\mathcal{I}$. Often, we will abbreviate $\Phi_{\bar{G}}^{\lambda,\gamma,\Gamma,X}(v)$ to $\Phi_{\bar{G}}^{\mathfrak{d},X}(v)$ or even $\Phi(v)$ if the subscript and superscript are clear from the context.

We have the following 'strong duality' result (proof in Appendix K.1):

Theorem J.1 (Strong Duality). *Let $\mathcal{I}$ be an instance of coordinated items .*

(a) Let $X \subset [0,H]$ be a discrete set. For any feasible solution $\mathfrak{p} = (a_{\bar{G}})$ of Equation 21 and any feasible solution $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ of the X-dual (Equation 22), it holds that:

$$\mathcal{P}(\mathfrak{p}) \le \mathscr{D}_X(\mathfrak{d}).$$

Equality holds if and only if the following conditions are satisfied almost everywhere:

$$\forall v : \Phi_{\bar{G}}^{\mathfrak{d}}(v) > 0 \implies a_{\bar{G}}(v) = 1. \tag{23a}$$

$$\forall v : \Phi_{\bar{G}}^{\mathfrak{d}}(v) < 0 \implies a_{\bar{G}}(v) = 0. \tag{23b}$$

$$\forall v : \lambda_{\bar{G}}(v) > 0 \implies a'_{\bar{G}}(v) = 0. \tag{23c}$$

$$\forall v : \gamma_G(v) > 0 \implies \int_0^v a_C(t)dt = \int_0^{g_G(v)} a_G(t)dt. \tag{23d}$$

$$\forall x \in X : \Gamma_G(x) > 0 \implies \int_0^x a_C(t)dt = \int_0^{g_G(x)} a_G(t)dt. \tag{23e}$$

(b) There exists a set X such that $\mathcal{P}(\mathcal{I}) = \mathscr{D}_X(\mathcal{I})$.

The main reason we provide a proof for this 'strong duality' result is that the variables are parametrized by a continuous variable. We could not find any results for such variables that subsume our setting. Our proof of Theorem J.1 works by showing, for all $\epsilon > 0$, a discrete linear program that has almost the same primal and dual value (up to terms depending on $\sqrt{\epsilon}$). Since strong duality holds for discrete systems, this gives us that the duality gap of our linear program is small. Theorem J.1 then follows as the feasible region is closed.

J.1.3. Our Framework Fix an instance $\mathcal{I}$ and discrete set $X \subseteq [0,H]$.

The Dual Framework. Let $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ be a feasible solution for the X-dual, *i.e.*, Equation 22. Define

$$\underline{r}_{\bar{G}}(\mathfrak{d}) = \inf\{v \mid \Phi_{\bar{G}}^{\mathfrak{d},X}(v) = 0\} \qquad \overline{r}_{\bar{G}}(\mathfrak{d}) = \sup\{v \mid \Phi_{\bar{G}}^{\mathfrak{d},X}(v) = 0\}.$$

If the inf (resp. sup) is over an empty set, we define it to be H (resp. 0).

We sometimes refer to γ and Γ as representing *flow, e.g.*, we say that there is flow from (C, v) to A if $\gamma_A(v) > 0$ or $\Gamma_A(v) > 0$.

An interval $[\underline{y}, \overline{y}] \subseteq [0, H]$ is said to be *ironed* on a item $\bar{G}$ if $\lambda_{\bar{G}}(\underline{y}) = \lambda_{\bar{G}}(\overline{y}) = 0$ and for all $v \in (\underline{y}, \overline{y})$, we have $\lambda_{\bar{G}}(v) > 0$.

The Primal Framework. Let $\mathfrak{p} = (a_{\bar{G}})$ be a feasible primal solution for Equation 21. Let $Y = [\underline{y}, \overline{y}] \subseteq [0, H]$ be an interval. Define $\mathsf{MC}_{\bar{G}}(Y, \mathfrak{p}) = |\{\alpha \mid \exists v \in Y : a_{\bar{G}}(v) = \alpha\}|$ to be the number of distinct values taken by the function $a_{\bar{G}}$ over the interval Y. Also, define $\mathsf{MC}(Y, \mathfrak{p}) = \max_{\bar{G} \in \mathsf{Items}} \mathsf{MC}_{\bar{G}}(Y, \mathfrak{p})$. We omit the argument Y if it is $[0, H]$.

We define the *menu complexity* of the instance $\mathcal{I}$, $\mathsf{MC}(\mathcal{I}) = \min_{\mathfrak{p}: \mathcal{P}(\mathfrak{p}) = \mathcal{P}(\mathcal{I})} \mathsf{MC}(\mathfrak{p})$ to be the smallest menu complexity of any optimal solution to Equation 21.

J.1.4. Formal Statements of our Results

Theorem J.2. *There exists an instance* $\mathtt{Uncountable}$ *of coordinated items such that* $\mathsf{MC}(\mathtt{Uncountable})$ *is uncountably infinite. Furthermore, the instance* $\mathtt{Uncountable}$ *satisfies* $g_B(v) = v$ *and all the distributions* $f_{\bar{G}}$ *are DMR*[15].

Theorem J.3. *For any instance* $\mathcal{I}$ *such that the functions* g_G *are piecewise linear, we have* $\mathsf{MC}(\mathcal{I})$ *is at most countably infinite.*

Theorem J.4. *There exists an instance* $\mathtt{Countable}_1$ *of coordinated items such that* $\mathsf{MC}(\mathtt{Countable}_1)$ *is countably infinite. Furthermore, the instance* $\mathtt{Countable}_1$ *satisfies* $g_B(v) = v$ *and all the distributions* $f_{\bar{G}}$ *are DMR, and the function* g_A *is piecewise linear.*

Theorem J.5. *There exists an instance* $\mathtt{Countable}_2$ *of coordinated items such that* $\mathsf{MC}(\mathtt{Countable}_2)$ *is countably infinite. Furthermore, the instance* $\mathtt{Countable}_2$ *satisfies* $g_B(v) = v$ *and the function* g_A *is piecewise linear with only* 2 *segments.*

J.2. Master Theorem

For our lower bounds, we will construct instances that have a large menu complexity. To show a lower bound on the menu complexity, we will define a feasible solution to the X-dual (Equation 22), for some X, show that it is optimal, and then show that any feasible primal that satisfies complementary slackness (Equation 23) with this dual must have a large menu complexity. Below, we prove a 'Master Theorem' (Theorem J.6) (proof in Appendix K.2) that constructs instances together with a feasible dual solution with some desirable properties.

Theorem J.6 (Master Theorem). *Let* $H > 0$ *be fixed. Suppose that, for all* $\bar{G} \in \mathsf{Items}$*, points* $\underline{\rho}_{\bar{G}} < \overline{\rho}_{\bar{G}} \in [1, H]$ *are given. Let* X_A*,* X_B *be discrete subsets of* $[\underline{\rho}_C, \overline{\rho}_C]$*. Consider finite or infinite sequences of disjoint intervals*

$$\mathcal{Y}_G = \{(\underline{y}_{G,i}, \overline{y}_{G,i})\}_{i \geq 0} \qquad \mathcal{Z}_G = \{(\underline{z}_{G,i}, \overline{z}_{G,i})\}_{i \geq 0}$$

where $\underline{y}_{G,i}, \overline{y}_{G,i} \in [\underline{\rho}_G, \overline{\rho}_G]$ *and* $\underline{z}_{G,i}, \overline{z}_{G,i} \in [\underline{\rho}_C, \overline{\rho}_C]$*. Then, for any invertible functions* $g'_G : [0, H] \to [0, H]$*, there exists an instance* $\mathcal{I} = (q, f_{\bar{G}}, g_G)$ *and an* $(X_A \cup X_B)$*-dual* $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ *such that the following hold:*

- $g_G(v) = g'_G(v)$ *for all* $v \in [0, H]$.
- *The value of* $\Phi^{\mathfrak{d}, X_A \cup X_B}_{\bar{G}}(v) f_{\bar{G}}(v)$ *is non-decreasing and* $\underline{r}_{\bar{G}}(\mathfrak{d}) = \underline{\rho}_{\bar{G}}$ *and* $\overline{r}_{\bar{G}}(\mathfrak{d}) = \overline{\rho}_{\bar{G}}$.
- $\lambda_C(v) = 0$ *throughout and* $\lambda_G(v) > 0 \iff v \in (\underline{y}_{G,i}, \overline{y}_{G,i})$ *for some* i.
- $\gamma_G(v) > 0 \iff v \in (\underline{z}_{G,i}, \overline{z}_{G,i})$ *for some* i.

[15] Recall that a distribution is *DMR* if the Myerson's virtual value function is non-decreasing.

- $\Gamma_G(v) > 0 \iff v \in X_G$ *for some* i.

Moreover, if $\mathcal{Y}_G$ *is empty, then the distributions* $f_{\bar{G}}$ *are DMR.*

J.3. Lower Bounds

J.3.1. DMR Distributions Let $\mathfrak{a} : [0,H] \to [0,1]$ be a non-decreasing function that is continuous except at countably many points. In this subsection, we construct an instance $\mathcal{I} = (q,f,g)$ of coordinated items such any optimal solution of Equation 21 for $\mathcal{I}$ satisfies that $a_A = \mathfrak{a}$[16]. In our construction, the distributions $f_{\bar{G}}$ are DMR and $g_B(v) = v$.

Two important instantiations of this general procedure prove Theorem J.2 and Theorem J.4. For Theorem J.2, we construct `Uncountable` by setting need $\mathfrak{a}(v) = v/H$. For Theorem J.4, we construct $\texttt{Countable}_1$ by setting $\mathfrak{a}$ to be a function that takes countably many values. A concrete example of such a function is one that has countably many "steps" as it moves from 0 to 1. We take care that the function g_A is piecewise linear in $\texttt{Countable}_1$.

The Instance For notational convenience, we work in the range $[0,H+3]$ in this subsubsection. Consider an increasing function $\mathfrak{a} : [0,H] \to [0,1]$ and let $\mathfrak{A}(v) = \int_0^v \mathfrak{a}(t)dt$. Let $\rho = H+2-\mathfrak{A}(H) \in [2,H+2]$. Note that $\mathfrak{A}^{-1}(v)$ is well defined for $v > 0$ and define:

$$g'_A(v) = \begin{cases} v & ,0 \le v \le 1 \\ \frac{v-1}{\rho-1} + 1 & ,1 < v \le \rho \\ \mathfrak{A}^{-1}(v-\rho) + 2 & ,\rho < v \le H+2 \\ v & ,H+2 < v \le H+3 \end{cases}.$$

It is readily seen seen that g'_A satisfies $\mathfrak{A}(H) - \mathfrak{A}\left(g'_A(v) - 2\right) = H - (v-2)$ in the interval $(\rho, H+2]$. Let $g'_B(v) = v$ and apply Theorem J.6 with g'_G and

- $\underline{\rho}_A = 2, \overline{\rho}_A = H+2, \underline{\rho}_B = \overline{\rho}_B = \rho, \underline{\rho}_C = 1, \overline{\rho}_C = H+2$.
- $X_G, \mathcal{Y}_G$ are empty
- $\mathcal{Z}_G(v) = \{(\rho, H+2)\}$.

This gives an instance $\mathcal{I} = (q,f,g)$ and an $(X_A \cup X_B)$-dual solution $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ such that the distributions $f_{\bar{G}}$ are DMR and :

- For all v, $g_G(v) = g'_G(v)$ and $\underline{r}_{\bar{G}} = \underline{\rho}_{\bar{G}}$ and $\overline{r}_{\bar{G}} = \overline{\rho}_{\bar{G}}$.
- $\lambda_{\bar{G}}(v), \Gamma_G(v)$ are 0 throughout.
- $\gamma_G(v) > 0 \iff v \in (\rho, H+2)$.

The Analysis Recall $\mathfrak{a}$, ρ, and $\mathcal{I}$ constructed above. Define a feasible primal solution $\mathfrak{p}^* = (a^*_{\bar{G}})$ of Equation 21 for $\mathcal{I}$ as:

$$a^*_A(v) = \begin{cases} 0 & ,0 \le v \le 2 \\ \mathfrak{a}(v-2) & ,2 < v \le H+2 \\ 1 & ,H+2 < v \le H+3 \end{cases} \qquad a^*_B(v) = a^*_C(v) = \begin{cases} 0 & ,0 \le v \le \rho \\ 1 & ,\rho < v \le H+3 \end{cases}$$

In Theorem J.7 (proof in Appendix K.3), we show that $\mathfrak{p}^*$ and $\mathfrak{d}$ satisfy complementary slackness. To finish our menu complexity lower bound, we argue that *any* primal $\mathfrak{p} = (a_{\bar{G}})$ that satisfies complementary slackness with $\mathfrak{d}$ must have $a_A = a^*_A$. This proof is in Theorem J.8.

[16] We abuse notation slightly here. What is meant is that a_A takes the same values over the interval $[2, H+2]$ as $\mathfrak{a}$ takes over the interval $[0, H]$ (see subsubsection J.3.1 for the exact statement).

Together with strong duality (Theorem J.1), Theorem J.7 shows that $\mathfrak{d}$ is optimal. Thus, any optimal primal for $\mathcal{I}$ must satisfy complementary slackness with $\mathfrak{d}$. Theorem J.8 says that $a_A = a^*_A$ for this primal and thus, it has a high menu complexity.

Lemma J.7. *The primal $\mathfrak{p}^*$ is feasible and $\mathfrak{p}^*$, $\mathfrak{d}$ satisfy complementary slackness (Equation 23).*

Lemma J.8. *Consider any feasible primal $\mathfrak{p} = (a_{\bar{G}})$ that satisfies complementary slackness with $\mathfrak{d}$ must $a_A = a^*_A$, where a^*_A is the allocation for item A in $\mathfrak{p}^*$.*

Proof. We reason from Equation 23. Specifically, the constraints (23a), (23b) for item B imply that $a_B(v) = 0$ for $v \in [0, \rho)$ and $a_B(v) = 1$ for $v \in (\rho, H+3]$. Thus, we have:

$$\int_0^{g_B(v)} a_B(t)dt = (v - \rho)\mathbb{1}_{v \geq \rho}.$$

Since $\gamma_G(v) > 0$ for all $v \in (\rho, H+2)$, we have by (23d) that $\int_0^{g_A(v)} a_A(t)dt = \int_0^{g_B(v)} a_B(t)dt = \int_0^{v} a_C(t)dt$ for all v in this range. Thus, for all $v \in (\rho, H+2)$:

$$\int_0^{g_A(v)} a_A(t)dt = v - \rho.$$

Since the right hand side in the equation above is independent of the primal, we get for all $v \in (\rho, H+2)$ that $\int_0^{g_A(v)} a_A(t)dt = \int_0^{g_A(v)} a^*_A(t)dt$. Thus, for all $v \in (2, H+2)$, we have $\int_0^{v} a_A(t)dt = \int_0^{v} a^*_A(t)dt$ implying $a_A(v) = a^*_A(v)$ in this range. The constraints (23a), (23b) for item A fix the allocation a_A outside $(2, H+2)$. Combining, we get that $a_A = a^*_A$.

□

J.3.2. Proof of Theorem J.5 In the last subsection, we showed that instances can have high menu complexity, even when all the distributions $f_{\bar{G}}$ are DMR. The reason for high menu complexity is the complexity in the functions g_G. We now show that if the distributions $f_{\bar{G}}$ are not required to be DMR, even 'simple' (*e.g.*, piecewise linear with only 2 segments) functions g_G can have countably infinite menu complexity. Two segments are required because of the arguments in Appendix F. This is tight due to our upper bounds in subsection J.4.

The Instance $\texttt{Countable}_2$ For notational convenience, we work in the range $[0, H+1]$ in this subsubsection. Define:

$$g'_A(v)^{17} = \begin{cases} v & ,0 \leq v \leq 1 \\ \frac{v+1}{2} & ,1 < v \leq 2H/3 + 1 \\ 2V - H - 1 & ,2H/3 + 1 < v \leq H + 1 \end{cases} \qquad g'_B(v) = v$$

We define points $x_1 = \frac{4H}{5} + 1$, $y_1 = \frac{3H}{4} + 1$, and $x_i = \frac{8}{3}\frac{H}{2^i} + 1$, $y_i = \frac{8}{5}\frac{H}{2^i} + 1$ for $i > 1$. Note that the sequence x_i converges to $x = 1$.

In order to construct $\texttt{Countable}_2$, we apply Theorem J.6 with g'_G and

- $\underline{\rho}_A = g'_A(x)$, $\overline{\rho}_A = g'_A(x_2)$, $\underline{\rho}_B = g'_B(x)$, $\overline{\rho}_B = g'_B(x_1)$, $\underline{\rho}_C = x$, $\overline{\rho}_C = x_1$.
- $\mathcal{Y}_A = \{(g'_A(x_{2i+2}), g'_A(x_{2i}))\}$ for some $i > 0$. $\mathcal{Y}_B = \{(g'_B(x_{2i+1}), g'_B(x_{2i-1}))\}$ for some $i > 0$.
- $\mathcal{Z}_G$ is empty and $X_G = \{y_i\}_{i>0}$.

This gives an instance $\texttt{Countable}_2 = (q, f, g)$ and an $(X_A \cup X_B)$-dual solution $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ such that :

[17] Note that this function is piecewise linear with 3 segments and not 2. However, the first segment is just for notational ease and can be removed.

- For all v, $g_G(v) = g'_G(v)$ and $\underline{r}_{\bar{G}} = \underline{\rho}_{\bar{G}}$ and $\overline{r}_{\bar{G}} = \overline{\rho}_{\bar{G}}$.
- $\lambda_C(v), \gamma_G(v)$ are 0 throughout. $\lambda_A(v) > 0$ if and only if $v \in (g_A(x_{2i+2}), g_A(x_{2i}))$ for some $i > 0$ and $\lambda_B(v) > 0$ if and only if $v \in (g_B(x_{2i+1}), g_B(x_{2i-1}))$ for some $i > 0$.
- $\forall i > 0 : \Gamma_G(y_i) > 0$.

The Analysis Recall the definitions of x_i, y_i and the instance $\texttt{Countable}_2$ above. Define a feasible primal solution $\mathfrak{p} = (a^*_{\bar{G}})$ of Equation 21 for $\texttt{Countable}_2$ as:

$$a^*_A(v) = \begin{cases} 0 & ,0 \le v \le g_A(x) \\ \frac{40}{13 \cdot 4^i} & ,g_A(x_{2i+2}) \le v < g_A(x_{2i}) \\ 1 & ,g_A(x_2) \le v \le H+1 \end{cases} \qquad a^*_B(v) = \begin{cases} 0 & ,0 \le v \le g_B(x) \\ \frac{40}{13 \cdot 4^i} & ,g_B(x_{2i+1}) \le v < g_B(x_{2i-1}) \\ 1 & ,g_B(x_1) \le v \le H+1 \end{cases}$$

$$a^*_C(v) = \begin{cases} 0 & ,0 \le v \le x \\ \frac{20}{13 \cdot 2^i} & ,y_{i+1} \le v < y_i \\ 1 & ,y_1 \le v \le H+1 \end{cases}$$

We proceed exactly as in subsubsection J.3.1. In Theorem J.9 (proof in Appendix K.4), we show that $\mathfrak{p}$ and $\mathfrak{d}$ satisfy complementary slackness. To finish our menu complexity lower bound, we argue that *any* primal that satisfies complementary slackness with $\mathfrak{d}$ must have infinite menu complexity.

Together with strong duality (Theorem J.1), Theorem J.9 shows that $\mathfrak{d}$ is optimal. Thus, any optimal primal for $\texttt{Countable}_2$ must satisfy complementary slackness with $\mathfrak{d}$. Theorem J.10 says that such a primal has infinite menu complexity

Lemma J.9. *There primal $\mathfrak{p}$ is feasible for $\texttt{Countable}_2$ and $\mathfrak{p}$,$\mathfrak{d}$ satisfy complementary slackness.*

Lemma J.10. *Any feasible primal $\mathfrak{p}$ for $\texttt{Countable}_2$ with a finite menu complexity does not satisfy complementary slackness with $\mathfrak{d}$.*

Proof. Proof by contradiction. Let $\mathfrak{p} = (a_{\bar{G}})$ be a feasible primal with finite menu complexity that satisfies complementary slackness with $\mathfrak{d}$.

Let i^* be the largest i such that $a_A(g_A(y_i)) > a_A(g_A(y_{i+1}))$ or $a_B(g_B(y_i)) > a_B(g_B(y_{i+1}))$. If no such i^* exists, we define $i^* = 0$. Since $\mathsf{MC}(\mathfrak{p})$ is assumed to be finite and the allocation is monotone (by constraint (21c)), i^* is well defined and $a_A(g_A(y_i))$ and $a_B(g_B(y_i))$ is constant for all $i > i^*$. Let the constant values be π_A and π_B respectively.

Since $\mathfrak{p}$, $\mathfrak{d}$ satisfy Equation 23, we have, by Equation 23c

$$a_G(v) = \pi_G \qquad ,\forall v \in (g_G(x), g_G(y_{i^*+1})] \tag{24}$$

Also, for all $i > 0$ we have by the constraint (23e) (applied once to both y_i and y_{i+1})

$$\int_{g_A(y_{i+1})}^{g_A(y_i)} a_A(t)dt = \int_{y_{i+1}}^{y_i} a_C(t)dt = \int_{g_B(y_{i+1})}^{g_B(y_i)} a_B(t)dt \tag{25}$$

We derive a contradiction in two steps. First, we prove

Claim. $\pi_A = 2\pi_B$.

Proof. Consider an i larger than $i^* + 10$. Using Equation 24 and Equation 25, we get

$$\begin{aligned} \pi_A \left(g_A(y_i) - g_A(y_{i+1})\right) &= \pi_B \left(g_B(y_i) - g_B(y_{i+1})\right) \\ \pi_A \left(y_i - y_{i+1}\right) &= 2\pi_B \left(y_i - y_{i+1}\right) && \text{(Definition of } g_A\text{)} \\ \pi_A &= 2\pi_B. \end{aligned}$$

□

Claim. $a_A(g_A(y_2)) \neq 2a_B(g_B(y_2))$.

Proof. Proof by contradiction. Suppose that $a_A(g_A(y_2)) = 2a_B(g_B(y_2)) = 2\pi$. By Equation 23a, we have $a_A(g_A(v)) = 1$ for all $v > x_2$. Since y_2, y_1 are in the same ironed interval on B and y_2, x_2 are in the same ironed interval on A, we have using Equation 25 that

$$g_A(y_1) - g_A(x_2) + 2\pi\left(g_A(x_2) - g_A(y_2)\right) = \pi\left(g_B(y_1) - g_B(y_2)\right).$$

Plugging in the values of y_1, x_2, y_2, we have

$$\frac{H}{2} - \frac{H}{3} + 2\pi\left(\frac{H}{3} - \frac{H}{5}\right) = \pi\left(\frac{3H}{4} - \frac{2H}{5}\right),$$

a contradiction to $\pi < 1$. □

These two claims together with Equation 24 establish that $i^* > 1$. We now give a contradiction assuming i^* is even. A similar argument works if i^* is odd. Since $g_A(y_{i^*}), g_A(y_{i^*+1})$ lie in the same ironed interval in A, we have $a_A(g_A(y_{i^*})) = a_A(g_A(y_{i^*+1})) = \pi_A$. By choice of i^*, we have $\pi'_B = a_B(g_B(y_{i^*})) > \pi_B$. By Equation 25, we have

$$\frac{\pi_A}{2}(y_{i^*} - y_{i^*+1}) = \pi_A(g_A(y_{i^*}) - g_A(y_{i^*+1})) = \pi'_B(y_{i^*} - x_{i^*+1}) + \pi_B(x_{i^*+1} - y_{i^*+1}) > \pi_B(y_{i^*} - y_{i^*+1}),$$

a contradiction to Equation J.3.2. □

J.4. Upper bounds

In this subsection, we prove that for any instance $\mathcal{I}$ such that the functions $g_A(\cdot)$ and $g_B(\cdot)$ are piecewise linear, we have $\mathsf{MC}(\mathcal{I})$ is at most countably infinite (Theorem J.3). This result is tight by our arguments in subsection J.3

Our line of argument is as follows: Fix an instance $\mathcal{I}$. By Theorem J.1, there exists an X, a primal solution $\mathfrak{p} = (a_{\bar{G}})$ and an X-dual solution $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ that satisfy complementary slackness. From $\mathfrak{p}, \mathfrak{d}$, we construct another primal solution $\hat{\mathfrak{p}}$ such that $\mathsf{MC}(\hat{\mathfrak{p}})$ is small and $\hat{\mathfrak{p}}$, $\mathfrak{d}$ satisfy complementary slackness. Thus, the primal $\hat{\mathfrak{p}}$ also defines an optimal revenue mechanism. The menu complexity of this new mechanism gives us our upper bound on $\mathsf{MC}(\mathcal{I})$.

We note that this technique is markedly different from that employed in Appendix F where we assume an optimal dual and describe a recovery algorithm that reads an optimal primal from the optimal dual. Here, we assume *both*, an optimal dual and an optimal primal[18], and prove that such a primal can be improved to have a lower menu complexity.

J.4.1. Splitting Our procedure to improve the primal makes extensive use of the following "splitting" operation:

Definition J.1. *Let $a : \mathbb{R} \to \mathbb{R}$ be a function and consider the interval $s = [x, y]$, where $x, y \in \mathbb{R}$. Define the function $a \circ s$ as :*

$$a \circ s(v) = \begin{cases} a(v) & , v \notin s \\ \frac{1}{y-x}\int_x^y a(t)dt & , v \in s \end{cases}$$

[18]Note that we need Theorem J.1 to assume that there exists an optimal primal-dual pair that satisfies complementary slackness.

The following is easily observed for all $v \notin s$:

$$\int_{-\infty}^{v} a \circ s(t)dt = \int_{-\infty}^{v} a(t)dt \tag{26}$$

If s_1 and s_2 are two disjoint intervals, then $a \circ s_1 \circ s_2 = a \circ s_2 \circ s_1$. We will use $a \circ s_1 s_2$ to denote this common value.

Remark 5. *Let a be a function and s be an interval. For any $x, y, z \in s$, we have*

$$\int_{-\infty}^{z} a \circ s(t)dt = \frac{1}{y-x}\left[(y-z)\int_{-\infty}^{x} a(t)dt + (z-x)\int_{-\infty}^{y} a(t)dt\right].$$

Lemma J.11. *Let a be an increasing function and s be an interval. We have $\int_{-\infty}^{z} a(t)dt \le \int_{-\infty}^{z} a \circ s(t)dt$. Moreover, equality holds if $v \notin s$ or a is constant over the interval s.*

Proof. Let $s = [x, y]$. Since a is increasing, we have

$$\int_{x}^{y} a(t)dt \ge \int_{x}^{z} a(t)dt + \frac{y-z}{z-x}\int_{x}^{z} a(t)dt = \frac{y-x}{z-x}\int_{x}^{z} a(t)dt.$$

Rearranging gives the result. The moreover part can be using Equation 26 and Definition J.1. □

J.4.2. Proof of Theorem J.3 Fix an instance $\mathcal{I}$ and let $\mathfrak{p}, \mathfrak{d}$, be an optimal primal dual pair ($\mathfrak{d}$ is X-dual for some X). Without loss of generality, we can assume that the product $\Phi_{\bar{G}}^{\mathfrak{d},X}(v) f_{\bar{G}}(v)$ is non-decreasing (see Appendix A). We define the function $\operatorname{sgn}(x)$ to be 1 if $x > 0$, 0 if $x = 0$, and -1 if $x < 0$. Using $\mathfrak{d}$, define the following notion of a strip.

Definition J.2 (Strip). *A strip s is an interval $[x, y]$, where $x, y \in [0, H]$, such that the following hold:*

- *For all $\bar{G}$, there functions $g_{\bar{G}}$ are linear over s*
- *For all $\bar{G}$, the function $\operatorname{sgn}\left(\Phi_{\bar{G}}^{\mathfrak{d},X}(g_{\bar{G}}(v)) f_{\bar{G}}(g_{\bar{G}}(v))\right)$ is constant for $v \in s$.*
- *For all $\bar{G}$, the function $\operatorname{sgn}\left(\lambda_{\bar{G}}(g_{\bar{G}}(v))\right)$ is constant for $v \in s$.*

The following holds for any strip.

Lemma J.12. *For $\mathcal{I}, \mathfrak{p}, \mathfrak{d}$ as above, let s be any strip. There exists another primal solution $\hat{\mathfrak{p}}$ of $\mathcal{I}$ such that*

- *$\hat{a}_{\bar{G}}(g_{\bar{G}}(v)) = a_{\bar{G}}(g_{\bar{G}}(v))$ for all $v \notin s$.*
- *$\hat{\mathfrak{p}}, \mathfrak{d}$ satisfy complementary slackness (Equation 23).*
- *For all $\bar{G}$, it holds that $\mathsf{MC}_{\bar{G}}\left([g_{\bar{G}}(x), g_{\bar{G}}(y)], \hat{a}_{\bar{G}}\right) \le 10$.*

The proof of this lemma spans the rest of this subsection.

Proof. Let $s = [x, y]$. Define the function $u_{\bar{G}}(v) = \int_0^{g_{\bar{G}}(v)} a_{\bar{G}}(t)dt$[19]. Since the functions $g_{\bar{G}}$ are continuous, and $a_{\bar{G}}$ is continuous except at countably many points, we have that $u_{\bar{G}}$ is also continuous. Define points $\underline{z}, \overline{z}$ as follows[20]:

$$\underline{z} = \inf_{v \in [x,y]} \{v \mid u_A(v) = u_B(v)\} \qquad \overline{z} = \sup_{v \in [x,y]} \{v \mid u_A(v) = u_B(v)\}.$$

[19] Note that this is just the utility of a bidder with type $(\bar{G}, g_{\bar{G}}(v))$.

[20] Throughout this subsection, we define several infimums and supremums. In case the argument to any of these is empty, we simply drop those terms from where they are used. For example, if $\underline{z}, \overline{z}$ are not defined, we simply use $a_G \circ [g_G(x), g_G(y)]$ instead of $a_G \circ [g_G(x), g_G(\underline{z})][g_G(\underline{z}), g_G(\overline{z})][g_G(\overline{z}), g_G(y)]$ below.

We use $\underline{z}$ and $\overline{z}$ to define:

$$\hat{a}_G = a_G \circ [g_G(x), g_G(\underline{z})][g_G(\underline{z}), g_G(\overline{z})][g_G(\overline{z}), g_G(y)].$$
$$\hat{u}_G(v) = \int_0^{g_G(v)} \hat{a}_G(t)dt.$$

Now, define:

$$z_1 = \inf_{v \in [x,\underline{z}]} \{v \mid \max(\hat{u}_A(v), \hat{u}_B(v)) = u_C(v)\} \qquad z_2 = \sup_{v \in [x,\underline{z}]} \{v \mid \max(\hat{u}_A(v), \hat{u}_B(v)) = u_C(v)\}.$$
$$z_3 = \inf_{v \in [\underline{z},\overline{z}]} \{v \mid \max(\hat{u}_A(v), \hat{u}_B(v)) = u_C(v)\} \qquad z_4 = \sup_{v \in [\underline{z},\overline{z}]} \{v \mid \max(\hat{u}_A(v), \hat{u}_B(v)) = u_C(v)\}.$$
$$z_5 = \inf_{v \in [\overline{z},y]} \{v \mid \max(\hat{u}_A(v), \hat{u}_B(v)) = u_C(v)\} \qquad z_6 = \sup_{v \in [\overline{z},y]} \{v \mid \max(\hat{u}_A(v), \hat{u}_B(v)) = u_C(v)\}.$$

Finally, we define:

$$\hat{a}_C = a_C \circ [x, z_1][z_1, z_2][z_2, \underline{z}][\underline{z}, z_3][z_3, z_4][z_4, \overline{z}][\overline{z}, z_5][z_5, z_6][z_6, y].$$
$$\hat{u}_C(v) = \int_0^{v} \hat{a}_C(t)dt.$$

Our primal $\hat{\mathfrak{p}}$ is defined by the allocations $(\hat{a}_{\bar{G}})$. Note that item 1 and item 3 are straightforward from Definition J.1. We only concentrate on item 2.

For item 2, we verify each of the constraints in Equation 23. For Equation 23a, observe that if $\Phi_{\bar{G}}(g_{\bar{G}}(v)) > 0$ for some $v \in s$ and some $\bar{G}$, then, since s is a strip, $\Phi_{\bar{G}}(g_{\bar{G}}(v)) > 0$ throughout s. Thus, the allocation $a_{\bar{G}}$ is 1 throughout s and our operations have no effect. If $v \notin s$, then the result follows as $\mathfrak{p}, \mathfrak{d}$ satisfied complementary slackness. A similar argument verifies Equation 23b and Equation 23c.

For Equation 23d, consider a v, G such that $\gamma_G(v) > 0$. If $v \notin s$, then the result follows because of 5 and the fact that $\mathfrak{p}, \mathfrak{d}$ satisfy complementary slackness. If $v \in s$, then, since $\mathfrak{p}, \mathfrak{d}$ satisfied complementary slackness, we have $u_G(v) = u_C(v)$. Thus, by Theorem J.11, $\hat{u}_G \geq u_C(v)$ implying that $v \in [z_1, z_2] \cup [z_3, z_4] \cup [z_5, z_6]$. Suppose that $v \in [z_1, z_2]$. The other cases are similar. We have:

$$\begin{aligned}
\hat{u}_G(v) &= \frac{\hat{u}_G(z_1)(g_G(z_2) - g_G(v)) + \hat{u}_G(z_2)(g_G(v) - g_G(z_1))}{g_G(z_2) - g_G(z_1)} && (5)\\
&= \frac{\hat{u}_G(z_1)(z_2 - v) + \hat{u}_G(z_2)(v - z_1)}{z_2 - z_1} && (g_G(v) = m_G v + c_G \text{ over a strip})\\
&= \frac{\hat{u}_C(z_1)(z_2 - v) + \hat{u}_C(z_2)(v - z_1)}{z_2 - z_1} && (\text{Definition of } z_1, z_2)\\
&= \hat{u}_C(v) && (5)
\end{aligned}$$

Equation 23e is verified similarly.

□

Proof of Theorem J.3. Since the functions g_G are assumed to be piecewise linear, we can partition the range $[0, H]$ into countably many disjoint strips. Theorem J.3 follows by applying Theorem J.12 to each of the strips. □

K. Missing Proofs From Appendix J

K.1. The Proof of Theorem J.1

K.1.1. The Proof of item (a)

Proof. Fix $X \subset [0,H]$. Observe that :

$$\mathscr{D}_X(\mathfrak{d}) = \sum_{\bar{G}} \int_0^H q_{\bar{G}} f_{\bar{G}}(t) \max\left(0, \Phi_{\bar{G}}(t)\right) dt \geq \sum_{\bar{G}} \int_0^H q_{\bar{G}} f_{\bar{G}}(t) \Phi_{\bar{G}}(t) a_{\bar{G}}(t) dt$$

Using the definition of Φ, we get

$$\mathscr{D}_X(\mathfrak{d}) \geq \mathcal{P}(\mathfrak{p}) + \sum_{\bar{G}} \int_0^H a_{\bar{G}}(t) \left(-\lambda'_{\bar{G}}(t) - \sum_{X \ni x > g_{\bar{G}}^{-1}(t)} \Gamma_{\bar{G}}(x) - \int_{g_{\bar{G}}^{-1}(t)}^H \gamma_{\bar{G}}(s) ds \right) dt$$

First, note that, integrating by parts, we have $\sum_{\bar{G}} \int_0^H a_{\bar{G}}(t) \lambda'_{\bar{G}}(t) dt = -\sum_{\bar{G}} \int_0^H a'_{\bar{G}}(t) \lambda_{\bar{G}}(t) dt \leq 0$ by Equation 21c. Fix any $v \in [0,H]$. Grouping all the terms with $\gamma_G(v)$, we observe that $\gamma_G(v)$ is multiplied by $\int_0^v a_C(t)dt - \int_0^{g_G(v)} a_G(t)dt \geq 0$ by Equation 21b. Similarly, fix any $x \in X$. Grouping all the terms with $\Gamma_G(x)$, we observe that $\Gamma_G(x)$ is multiplied by $\int_0^x a_C(t)dt - \int_0^{g_G(x)} a_G(t)dt \geq 0$ by Equation 21b.

Thus, we have $\mathscr{D}_X(\mathfrak{d}) \geq \mathcal{P}(\mathfrak{p})$. To finish the proof, observe that the conditions in Equation 23 are exactly those needed to make these inequalities tight. □

K.1.2. The Proof of item (b)

Proof. We prove that for all $\epsilon > 0$, there is a set X_ϵ such that $\mathcal{P}(\mathcal{I}) \geq \mathscr{D}_{X_\epsilon}(\mathcal{I}) - \epsilon$. The statement then follows because the the union of the region in Equation 22 for all possible sets X is closed. Recall that the functions g_G are L-Lipschitz.

Fix $\epsilon > 0$ and let $\delta > 0$ be sufficiently small. Our proof proceeds by defining a *discrete* linear program (Equation 27) and its dual (Equation 28) for $\mathcal{I}, \delta$. Let the optimal value of Equation 27 be $\mathcal{P}_\delta(\mathcal{I})$ and the optimal value of Equation 28 be $\mathscr{D}_\delta(\mathcal{I})$. Since strong duality holds for discrete linear programs, we have $\mathcal{P}_\delta(\mathcal{I}) = \mathscr{D}_\delta(\mathcal{I})$. We also ensure that $\mathscr{D}_\delta(\mathcal{I}) = \mathcal{P}_\delta(\mathcal{I}) \leq H$ for all δ.

In Theorem K.1, we show that $\mathcal{P}(\mathcal{I}) \geq (1 - \sqrt{\delta})\mathcal{P}_\delta(\mathcal{I}) - L\sqrt{\delta}$. In Theorem K.2, we show that there is a set X_δ such that $\mathscr{D}_{X_\delta}(\mathcal{I}) \leq \mathscr{D}_\delta(\mathcal{I})$. Combining, we get

$$\mathcal{P}(\mathcal{I}) \geq (1-\sqrt{\delta})\mathcal{P}_\delta(\mathcal{I}) - L\sqrt{\delta} = (1-\sqrt{\delta})\mathscr{D}_\delta(\mathcal{I}) - L\sqrt{\delta} \geq (1-\sqrt{\delta})\mathscr{D}_{X_\delta}(\mathcal{I}) - L\sqrt{\delta} > \mathscr{D}_{X_\delta}(\mathcal{I}) - \epsilon,$$

for small enough δ. □

The rest of this subsection is devoted to defining and analyzing the discrete linear program, in order to prove Theorem K.1 and Theorem K.2.

K.1.3. The Discrete Linear Program We describe a discrete linear program for the instance $\mathcal{I}$. The instance $\mathcal{I}$ is fixed for the rest of this subsection. Without loss of generality, let $\delta > 0$ be such that $H/\delta = H'$ is an integer.

The Primal Consider the following optimization problem with the variables $a_{\bar{G}}(i)$ and $p_{\bar{G}}(i)$ for $0 \leq i \leq H'$. In this subsection, we abuse notation and write $g_{\bar{G}}(i)$ instead of $g_{\bar{G}}(i\delta)$. Define $\int_{g_{\bar{G}}(i-1)}^{g_{\bar{G}}(i)} f_{\bar{G}}(v)dv = \hat{f}_{\bar{G}}(i)$ and $\hat{f}_{\bar{G}}(0) = 0$.

$$\text{maximize} \mathcal{P}_\delta(a) = \sum_{\bar{G}} \sum_{i=1}^{H'} q_{\bar{G}} \hat{f}_{\bar{G}}(i) p_{\bar{G}}(i) \tag{27a}$$

subject to

$$g_C(i)a_C(i) - p_C(i) \geq g_G(i)a_G(i) - p_G(i) \quad ,\forall i \tag{27b}$$

$$g_{\bar{G}}(i)a_{\bar{G}}(i) - p_{\bar{G}}(i) \geq g_{\bar{G}}(i)a_{\bar{G}}(i+1) - p_{\bar{G}}(i+1) \quad ,\forall i \tag{27c}$$

$$g_{\bar{G}}(i)a_{\bar{G}}(i) - p_{\bar{G}}(i) \geq g_{\bar{G}}(i)a_{\bar{G}}(i-1) - p_{\bar{G}}(i-1) \quad ,\forall i \tag{27d}$$

$$p_{\bar{G}}(0) = a_{\bar{G}}(0) = 0 \tag{27e}$$

$$a_{\bar{G}}(i) \in [0,1] \quad ,\forall i \tag{27f}$$

It is easy to see why $\mathcal{P}_\delta(\mathcal{I}) \leq H$. We also have:

Theorem K.1. *The optimal value $\mathcal{P}_\delta(\mathcal{I})$ of Equation 27 satisfies $(1-\sqrt{\delta})\mathcal{P}_\delta(\mathcal{I}) \leq \mathcal{P}(\mathcal{I}) + L\sqrt{\delta}$.*

Proof. Let $\eta = \sqrt{\delta}$ and $(a_{\bar{G}}, p_{\bar{G}})$ be the optimal solution for Equation 27. Consider the set $\mathcal{T} = \{(G_j, a_j, p_j) \mid \exists i, G_j : a_{G_j}(i) = a_j, (1-\eta)p_{G_j}(i) = p_j\}$. Using the set $\mathcal{T}$, we now define a mechanism M for the continuous revenue optimization problem. Consider a type $\mathrm{t} = (\bar{G}, v) \in$ Types. Define:

$$(G_\mathrm{t}, a_\mathrm{t}, p_\mathrm{t}) = \arg\max_{(G_j,a_j,p_j)\in\mathcal{T}} \mathfrak{v}(\mathrm{t}, G_j)a_j - p_j.$$

Let M be the mechanism that allocates item G_t with probability a_t and charges price p_t to a bidder who reports t[21]. Observe that M is a truthful mechanism. By a standard argument, there exists a feasible solution $\mathfrak{p}$ to Equation 21 such that $\mathcal{P}(\mathcal{I}) \geq \mathcal{P}(\mathfrak{p}) = \mathbb{E}_{\mathrm{t}\sim(q,f)}[p_\mathrm{t}]$ is the expected revenue of M. We now prove that $\mathbb{E}_{\mathrm{t}\sim(q,f)}[p_\mathrm{t}] \geq (1-\eta)\mathcal{P}_\delta(\mathcal{I}) - L\eta$.

Couple a bidder in M with (continuous) type $\mathrm{t} = (\bar{G}, v)$ with a bidder with (discrete) type $(\bar{G}, i)$ where i is the smallest value such that $g_{\bar{G}}(i) > v$. The coupling is valid as $\int_{g_{\bar{G}}(i-1)}^{g_{\bar{G}}(i)} f_{\bar{G}}(v)dv = \hat{f}_{\bar{G}}(i)$. We show that $p_\mathrm{t} \geq (1-\eta)p_{\bar{G}}(i) - L\eta$. Taking the expectation on both sides gives the result.

Observe that:

$$\mathfrak{v}(\mathrm{t}, G_\mathrm{t})a_\mathrm{t} - p_\mathrm{t} \geq v a_{\bar{G}}(i) - (1-\eta)p_{\bar{G}}(i)$$

$$g_{\bar{G}}(i)a_{\bar{G}}(i) - p_{\bar{G}}(i) \geq \mathfrak{v}(\bar{G}, g_{\bar{G}}(i), G_\mathrm{t})a_\mathrm{t} - \frac{p_\mathrm{t}}{1-\eta}$$

Adding, we get that

$$\mathfrak{v}(\bar{G}, v, G_\mathrm{t})a_\mathrm{t} + g_{\bar{G}}(i)a_{\bar{G}}(i) + \frac{\eta p_\mathrm{t}}{1-\eta} \geq v a_{\bar{G}}(i) + \eta p_{\bar{G}}(i) + \mathfrak{v}(\bar{G}, g_{\bar{G}}(i), G_\mathrm{t})a_\mathrm{t}$$

Since $a_\mathrm{t}, a_{\bar{G}}(i) \in [0,1]$, we have

$$\begin{aligned}
p_\mathrm{t} &\geq (1-\eta)p_{\bar{G}}(i) + \frac{1-\eta}{\eta}\left((v - g_{\bar{G}}(i))a_{\bar{G}}(i) + \left(\mathfrak{v}(\bar{G}, g_{\bar{G}}(i), G_\mathrm{t}) - \mathfrak{v}(\bar{G}, v, G_\mathrm{t})\right)a_\mathrm{t}\right) \\
&\geq (1-\eta)p_{\bar{G}}(i) + \frac{1}{\eta}\left(g_{\bar{G}}(i-1) - g_{\bar{G}}(i)\right)a_{\bar{G}}(i) && (g_{\bar{G}}(i) > v \geq g_{\bar{G}}(i-1)) \\
&\geq (1-\eta)p_{\bar{G}}(i) - L\eta && (g_{\bar{G}} \text{ is } L\text{-Lipschitz})
\end{aligned}$$

□

The Dual Consider the following Lagrangian relaxation of Equation 27, where we Lagrangify the constraints Equation 27b using the variables $\Gamma_G(i)$, the constraints Equation 27c using the variables $\lambda^+_G(i)$, and the constraints Equation 27d using the variables $\lambda^-_G(i)$. We use the convention that $\Gamma_C(i) = -\Gamma_A(i) - \Gamma_B(i)$

$$\text{minimize} \qquad \mathscr{D}_\delta(\mathcal{I})(\lambda, \Gamma) = \sum_{\bar{G}} \sum_{i=1}^{H'} q_{\bar{G}} \hat{f}_{\bar{G}}(i) \max\left(0, \Phi^{\lambda,\Gamma}_{\bar{G}}(i)\right) \tag{28a}$$

subject to

$$\lambda^+_{\bar{G}}(i) + \lambda^-_{\bar{G}}(i) - \Gamma_{\bar{G}}(i) = q_{\bar{G}}\hat{f}_{\bar{G}}(i) + \lambda^+_{\bar{G}}(i-1) + \lambda^-_{\bar{G}}(i+1) \quad ,\forall i \tag{28b}$$

[21] We slightly deviate from our definition of a mechanism and allow G_t to be different from $\bar{G}$.

$$\lambda^+_{\bar{G}}(i), \lambda^-_{\bar{G}}(i) \geq 0 \qquad ,\forall i \tag{28c}$$
$$\gamma_G(i) \geq 0 \qquad ,\forall i \tag{28d}$$

where terms like $\lambda^+ \bar{G}(H+1)$ are defined to be 0 and

$$\Phi^{\lambda,\gamma}_{\bar{G}}(i) = \left(g_{\bar{G}}(i) - \frac{1}{q_{\bar{G}} \hat{f}_{\bar{G}}(i)} \left(\lambda^+_{\bar{G}}(i-1) \left(g_{\bar{G}}(i-1) - g_{\bar{G}}(i) \right) + \lambda^-_{\bar{G}}(i+1) \left(g_{\bar{G}}(i+1) - g_{\bar{G}}(i) \right) \right) \right)$$

Theorem K.2. *The optimal value $\mathscr{D}_\delta(\mathcal{I})$ of Equation 28 satisfies $\mathscr{D}_\delta(\mathcal{I}) \geq \mathscr{D}_X(\mathcal{I})$ where $X = \{i\delta \mid i \in \mathbb{Z} \cap [0, H']\}$.*

Proof. We proceed by defining a feasible solution of Equation 22 from the optimal solution of Equation 28 with the same value. Throughout this proof, we denote the variables of the discrete linear program using a 'hat'. Let $(\hat{\lambda}^+_{\bar{G}}, \hat{\lambda}^-_{\bar{G}}, \hat{\Gamma}_G)$ be the optimal solution to Equation 28. Define $\hat{F}_{\bar{G}}(v) = \sum_{i: g_{\bar{G}}(i) < v} \hat{f}_{\bar{G}}(i)$. Define an X-dual solution $\mathfrak{d} = (\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ for $\mathcal{I}$ as follows

$$\lambda_{\bar{G}}(v) = v q_{\bar{G}}(F_{\bar{G}}(v) - \hat{F}_{\bar{G}}(v)) + (v - g_{\bar{G}}(i))\lambda^+_{\bar{G}}(i) + (g_{\bar{G}}(i+1) - v)\frac{g_{\bar{G}}(i) - g_{\bar{G}}(i-1)}{g_{\bar{G}}(i+1) - g_{\bar{G}}(i)}\lambda^+_{\bar{G}}(i-1)$$
$$\gamma_G(v) = 0$$
$$\Gamma_G(i\delta) = \hat{\Gamma}_G(i)$$

where i is the largest integer such that $g_{\bar{G}}(i) < v$. Observe that $\mathfrak{d}$ is feasible and

$$q_{\bar{G}} f_{\bar{G}}(v) \Phi^{\mathfrak{d},X}_{\bar{G}}(v) = -q_{\bar{G}} \frac{g_{\bar{G}}(i+1)\left(1 - \hat{F}_{\bar{G}}(v)\right) - g_{\bar{G}}(i)\left(1 + \hat{f}_{\bar{G}}(i) - \hat{F}_{\bar{G}}(v)\right)}{g_{\bar{G}}(i+1) - g_{\bar{G}}(i)}$$
$$- \lambda^+_{\bar{G}}(i) + \frac{g_{\bar{G}}(i) - g_{\bar{G}}(i-1)}{g_{\bar{G}}(i+1) - g_{\bar{G}}(i)}\lambda^+_{\bar{G}}(i-1) - \sum_{j=i+1}^{H/\epsilon} \Gamma_{\bar{G}}(j)$$

Since $\Phi^{\mathfrak{d},X}_{\bar{G}}(v)$ is constant in $[g_{\bar{G}}(i), g_{\bar{G}}(i+1)]$, we have:

$$\int_{g_{\bar{G}}(i)}^{g_{\bar{G}}(i+1)} q_{\bar{G}} f_{\bar{G}}(t) \max\left(0, \Phi^{\mathfrak{d},X}_{\bar{G}}(t)\right) dt = \max\left(0, \int_{g_{\bar{G}}(i)}^{g_{\bar{G}}(i+1)} q_{\bar{G}} f_{\bar{G}}(t) \Phi^{\mathfrak{d},X}_{\bar{G}}(t) dt\right)$$

But,

$$\int_{g_{\bar{G}}(i)}^{g_{\bar{G}}(i+1)} q_{\bar{G}} f_{\bar{G}}(t) \Phi^{\mathfrak{d},X}_{\bar{G}}(t) dt$$
$$= g_{\bar{G}}(i) q_{\bar{G}} \hat{f}_{\bar{G}}(i) - \lambda^+_{\bar{G}}(i-1)\left(g_{\bar{G}}(i-1) - g_{\bar{G}}(i)\right) - \int_{g_{\bar{G}}(i)}^{g_{\bar{G}}(i+1)} \left(\lambda^+_{\bar{G}}(i) + q_{\bar{G}}\left(1 - \hat{F}_{\bar{G}}(t)\right) + \sum_{j=i+1}^{H'} \Gamma_{\bar{G}}(j) \right) dt$$
$$= g_{\bar{G}}(i) q_{\bar{G}} \hat{f}_{\bar{G}}(i) - \lambda^+_{\bar{G}}(i-1)\left(g_{\bar{G}}(i-1) - g_{\bar{G}}(i)\right) - \lambda^-_{\bar{G}}(i+1)\left(g_{\bar{G}}(i+1) - g_{\bar{G}}(i)\right) \quad \text{(Equation 28b)}$$
$$= q_{\bar{G}} \hat{f}_{\bar{G}}(i) \hat{\Phi}^{\hat{\lambda},\hat{\Gamma}}_{\bar{G}}(i)$$

Thus, $\int_{g_{\bar{G}}(i)}^{g_{\bar{G}}(i+1)} q_{\bar{G}} f_{\bar{G}}(t) \max\left(0, \Phi^{\mathfrak{d},X}_{\bar{G}}(t)\right) dt = q_{\bar{G}} \hat{f}_{\bar{G}}(i) \max\left(0, \hat{\Phi}^{\hat{\lambda},\hat{\Gamma}}_{\bar{G}}(i)\right)$. Summing over $i, \bar{G}$, we get that $\mathscr{D}_\delta(\mathcal{I}) = \mathscr{D}_X(\mathfrak{d}) \geq \mathscr{D}_X(\mathcal{I})$. □

K.2. Proof of Theorem J.6

We need the following technical lemma:

Lemma K.3. *Let $H > 1$ and $R(x) : [1,H] \to \mathbb{R}$ be a function such that $R(1) = 1$ and for all $x,y \in [1,H]$ such that $x < y$, we have $-\frac{y-x}{2H} < R(y) - R(x) < \frac{y-x}{2H}$. Then, there exists a distribution F supported on $[0,H]$ such that*

$$x(1-F(x)) = \begin{cases} x & ,0 \le x < 1 \\ R(x) & ,1 \le x \le H \end{cases}.$$

Proof. Define:

$$F(x) = \begin{cases} 0 & ,0 \le x < 1 \\ 1 - \frac{R(x)}{x} & ,1 \le x < H \\ 1 & ,x = H \end{cases}.$$

Observe that F satisfies the requirements of the theorem. To prove that F is a valid distribution, it is sufficient to show $F(x) < F(y)$ for all $x < y$. We first note that $1 - \frac{R(x)}{x} \in [0,1]$. Thus, the only case left is when $x,y \in [1,H]$. In this case,

$$\begin{aligned} xy\,(F(y) - F(x)) &= yR(x) - xR(y) = yR(x) - xR(x) + x(R(x) - R(y)) \\ &> yR(x) - xR(x) - x\frac{y-x}{2H} \\ &= (y-x)\left(R(x) - \frac{x}{2H}\right) = (y-x)\left(R(x) - 1 + \frac{2H-x}{2H}\right) \\ &> (y-x)\left(R(x) - R(1) + \frac{x-1}{2H}\right) > 0. \end{aligned}$$

□

Proof of Theorem J.6. Let $\mathcal{I} = (q, f_{\bar{G}}, g_G)$ be such that $q_{\bar{G}} = \frac{1}{3}$ and $g_G(v) = g'_G(v)$ for all v. In order to define $f_{\bar{G}}$, define

$$Q_{\bar{G}}(v) = \begin{cases} 1 + \frac{v-1}{100H} & ,1 \le v < \underline{\rho}_G \\ 1 + \frac{\underline{\rho}_G{}^{-1}}{100H} & ,\underline{\rho}_G \le v < \overline{\rho}_G \\ 1 + \frac{\overline{\rho}_G + \underline{\rho}_G{}^{-1} - v}{100H} & ,\overline{\rho}_G \le v \le H \end{cases}$$

$$\Lambda_G(v) = \frac{1}{100H^2} \sum_i \min\left((v - \underline{y}_{G,i})(v - \overline{y}_{G,i}), 0\right)$$

Let $\Lambda_C(v) = 0$ throughout. Also, define:

$$Z_{G,i}(v) = \begin{cases} \frac{1}{4^i} & ,v < \underline{z}_{G,i} \\ \frac{1}{4^i}\frac{\overline{z}_{G,i} - v}{\overline{z}_{G,i} - \underline{z}_{G,i}} & ,v \in [\underline{z}_{G,i}, \overline{z}_{G,i}] \\ 0 & ,v > \overline{z}_{G,i} \end{cases} \qquad Y_{G,i}(v) = \begin{cases} \frac{1}{4^i} & ,v < \underline{x}_{G,i} \\ 0 & ,v > \overline{x}_{G,i} \end{cases}$$

$$P_{1,G}(v) = \frac{1}{100H} \sum_i Z_{G,i}(v) \qquad P_{2,G}(v) = \frac{1}{100H} \sum_i Y_{G,i}(v)$$

Finally, define $P_{1,C}(v) = -P_{1,A}(v) - P_{1,B}(v)$ and $P_{2,C}(v) = -P_{2,A}(v) - P_{2,B}(v)$ and ,

$$R_{\bar{G}}(v) = Q_{\bar{G}}(v) + \frac{1}{q_{\bar{G}}}\left[\Lambda_{\bar{G}}(v) - \int_1^v P_{1,\bar{G}}(g_{\bar{G}}^{-1}(t))dt - \int_1^v P_{2,\bar{G}}(g_{\bar{G}}^{-1}(t))dt\right]$$

Observe that the function $R_{\bar{G}}$ satisfies all the requirements of Theorem K.3. Thus, there exist distributions $f_{\bar{G}}$ such that

$$x(1 - F_{\bar{G}}(x)) = \begin{cases} x & ,0 \le x < 1 \\ R_{\bar{G}}(x) & ,1 \le x \le H \end{cases}.$$

Set $\lambda_{\bar{G}} = -\Lambda_{\bar{G}}$, $\gamma_G(v) = -P'_{1,G}(v)$ and $\Gamma_G(\cdot) : X_G \to \mathbb{R}$ to be the unique function such that $\sum_{x>v} \Gamma(x) = P_{2,G}(v)$.

Observe that the dual defined by $(\lambda_{\bar{G}}, \gamma_G, \Gamma_G)$ satisfies all the requirements of the theorem.

□

K.3. Omitted Proofs in subsubsection J.3.1

Proof of Theorem J.7. We verify each of the constraints in Equation 23 and leave verifying feasibility using Equation 21 to the reader. The constraints (23a), (23b) are verified easily. The constraint (23c) is true because $\lambda_G(v) = 0$ throughout. The constraint (23d) holds because for all $v \in (\rho, H+2)$, we have

$$\int_0^v a^*_C(t)dt = v - \rho$$

$$\int_0^{g_B(v)} a^*_B(t)dt = \int_0^v a^*_B(t)dt = v - \rho$$

$$\int_0^{g_A(v)} a^*_A(t)dt = \int_2^{g_A(v)} \mathfrak{a}(t-2)dt = \int_0^{g_A(v)-2} \mathfrak{a}(t)dt = \mathfrak{A}(g_A(v) - 2) = v - \rho.$$

and the three quantities are equal. Finally, the constraint (23e) is satisfied because Γ is zero throughout.

□

K.4. Omitted Proofs in subsubsection J.3.2

Recall that $x_1 = \frac{4H}{5} + 1$, $y_1 = \frac{3H}{4} + 1$, and $x_i = \frac{8}{3}\frac{H}{2^i} + 1$, $y_i = \frac{8}{5}\frac{H}{2^i} + 1$ for $i > 1$. This implies the equations

$$\forall i > 1 : x_i - y_i = \frac{16H}{15 \cdot 2^i}. \tag{29a}$$

$$\forall i > 1 : y_i - x_{i+1} = \frac{4H}{15 \cdot 2^i}. \tag{29b}$$

$$\forall i > 1 : \frac{1}{2}(x_{i+1} - y_{i+1}) + 2(y_i - x_{i+1}) = (y_i - y_{i+1}) \tag{29c}$$

$$5y_2 + 16y_1 = 21x_2. \tag{29d}$$

We will need the following lemma:

Lemma K.4. *For all $i > 0$ and $v \in [y_{i+1}, y_i]$,*

$$\int_{g_G(y_{i+1})}^{g_G(v)} a^*_G(x)dx \le \int_{y_{i+1}}^{v} a^*_C(x)dx.$$

Moreover, equality holds if $v = y_i$.

Proof. We calculate the three quantities:

$$\int_{g_A(y_{i+1})}^{g_A(v)} a_A^*(t)dt = \begin{cases} \frac{10}{13}\left(g_A(v) - g_A(y_2)\right) & ,i = 1, v \le x_2 \\ \frac{10}{13}\left(g_A(x_2) - g_A(y_2)\right) + \left(g_A(v) - g_A(x_2)\right) & ,i = 1, v > x_2 \\ \frac{40}{13\cdot 2^i}\left(g_A(v) - g_A(y_{i+1})\right) & ,i \text{ is even} \\ \frac{40}{13\cdot 2^{i+1}}\left(g_A(v) - g_A(y_{i+1})\right) & ,i \text{ is odd}, v \le x_{i+1} \\ \frac{40}{13\cdot 2^{i+1}}\left(g_A(x_{i+1}) - g_A(y_{i+1})\right) + \frac{40}{13\cdot 2^{i-1}}\left(g_A(v) - g_A(x_{i+1})\right) & ,i \text{ is odd}, v > x_{i+1} \end{cases}$$

$$\int_{g_B(y_{i+1})}^{g_B(v)} a_B^*(t)dt = \begin{cases} \frac{40}{13\cdot 2^{i+2}}\left(v - y_{i+1}\right) & ,i \text{ is even}, v \le x_{i+1} \\ \frac{40}{13\cdot 2^{i+2}}\left(x_{i+1} - y_{i+1}\right) + \frac{40}{13\cdot 2^i}\left(v - x_{i+1}\right) & ,i \text{ is even}, v > x_{i+1} \\ \frac{40}{13\cdot 2^{i+1}}\left(v - y_{i+1}\right) & ,i \text{ is odd} \end{cases}$$

$$\int_{y_{i+1}}^{v} a_C^*(t)dt = \frac{20}{13\cdot 2^i}\left(v - y_{i+1}\right)$$

We now prove the result for $i = 1$. In this case, as the expressions for B and C are the same, it is sufficient to show that

$$\begin{aligned} \int_{g_A(y_{i+1})}^{g_A(v)} a_A^*(t)dt &= \begin{cases} \frac{5}{13}\left(v - y_2\right) & ,v \le x_2 \\ \frac{5}{13}\left(x_2 - y_2\right) + 2v - 2x_2 & ,v > x_2 \end{cases} \\ &= \begin{cases} \frac{5}{13}\left(v - y_2\right) & ,v \le x_2 \\ 2v - \frac{10y_2}{13} - \frac{16y_1}{13} & ,v > x_2 \end{cases} && \text{(Equation 29d)} \\ &\le \frac{10}{13}\left(v - y_2\right) = \int_{y_{i+1}}^{v} a_C^*(t)dt && \text{(As } v < y_1\text{)} \end{aligned}$$

We now prove for even $i > 1$. The case for odd i is similar. Note that $g_A(x_i) = \frac{x_i+1}{2}$ and $g_A(y_i) = \frac{y_i+1}{2}$ for all $i > 1$. In this case, as the expressions for A and C are the same, it is sufficient to show that

$$\begin{aligned} \int_{g_B(y_{i+1})}^{g_B(v)} a_B^*(t)dt &= \begin{cases} \frac{40}{13\cdot 2^{i+2}}\left(v - y_{i+1}\right) & ,v \le x_{i+1} \\ \frac{40}{13\cdot 2^{i+2}}\left(x_{i+1} - y_{i+1}\right) + \frac{40}{13\cdot 2^i}\left(v - x_{i+1}\right) & ,v > x_{i+1} \end{cases} \\ &= \begin{cases} \frac{10}{13\cdot 2^i}\left(v - y_{i+1}\right) & ,v \le x_{i+1} \\ \frac{20}{13\cdot 2^i}\left(y_i - y_{i+1}\right) + \frac{40}{13\cdot 2^i}\left(v - y_i\right) & ,v > x_{i+1} \end{cases} && \text{(Equation 29c)} \\ &\le \frac{20}{13}\left(v - y_i\right) = \int_{y_{i+1}}^{v} a_C^*(t)dt && \text{(As } v < y_1\text{)} \end{aligned}$$

The moreover part can be observed by putting $v = y_i$ in our equations. □

Proof of Theorem J.9.

Feasibility We verify the feasibility constraints in Equation 21. It is straightforward to verify that $a_{\bar{G}}^*$ are monotone (constraint (21c)) and take values in $[0, 1]$ (constraint (21d)). Finally, the constraint (21b) holds because of Theorem K.4.

Optimality We verify all the constraints in Equation 23. The constraints (23a), (23b) are straightforward to verify from the definition of a_G^*. The constraint (23c) follows from the fact that $\lambda_C(v)$ is 0 throughout and

$$\lambda_A(v) > 0 \implies \exists i > 0 : v \in \left(g_A(x_{2i+1}), g_A(x_{2i-1})\right) \implies a_A'^*(v) = 0$$
$$\lambda_B(v) > 0 \implies \exists i > 0 : v \in \left(g_B(x_{2i+2}), g_B(x_{2i})\right) \implies a_B'^*(v) = 0$$

The constraint (23d) holds because $\gamma_G(v)$ is 0 throughout. For the constraint (23e), we need to prove that for all y_i

$$\int_0^{g_A(y_i)} a_A^*(x)dx = \int_0^{g_B(y_i)} a_B^*(x)dx = \int_0^{y_i} a_C^*(x)dx$$

which holds because of (the moreover part of) Theorem K.4

□