

# The Optimality of Upgrade Pricing

Dirk Bergemann<sup>1</sup>, Alessandro Bonatti<sup>2</sup>, Andreas Haupt<sup>3</sup>, and Alex Smolin<sup>4</sup>[0000–0003–4740–2376]

<sup>1</sup> Yale University, Department of Economics, New Haven, USA  
`dirk.bergemann@yale.edu`

<sup>2</sup> Massachusetts Institute of Technology, Sloan School of Management, Cambridge, USA  
`bonatti@mit.edu`

<sup>3</sup> Massachusetts Institute of Technology, Institute for Data, Systems, and Society, Cambridge, USA  
`haupt@mit.edu`

<sup>4</sup> Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France  
`alexey.v.smolin@gmail.com`

**Abstract.** We consider a multiproduct monopoly pricing model. We provide sufficient conditions under which the optimal mechanism can be implemented via upgrade pricing—a menu of product bundles that are nested in the strong set order. Our approach exploits duality methods to identify conditions on the distribution of consumer types under which (a) each product is purchased by the same set of buyers as under separate monopoly pricing (though the transfers can be different), and (b) these sets are nested.

We exhibit two distinct sets of sufficient conditions. The first set of conditions weakens the monotonicity requirement of types and virtual values but maintains a *regularity* assumption, i.e., that the product-by-product revenue curves are single-peaked. The second set of conditions establishes the optimality of upgrade pricing for type spaces with *monotone marginal rates of substitution (MRS)*—the relative preference ratios for any two products are monotone across types. The monotone MRS condition allows us to relax the earlier regularity assumption.

Under both sets of conditions, we fully characterize the product bundles and prices that form the optimal upgrade pricing menu. Finally, we show that, if the consumer’s types are monotone, the seller can equivalently post a vector of single-item prices: upgrade pricing and separate pricing are equivalent.

**Keywords:** Revenue maximization; mechanism design; strong duality; upgrade pricing.

## 1 Introduction

### 1.1 Motivation and Results

Pricing multiple goods with market power is a canonical problem in the theory of mechanism design. It is also a challenge of growing importance and complexity for online retailers and service providers, such as Amazon and Netflix. Both in theory and in practice, designing the optimal *mixed bundling* mechanism, (i.e., pricing every subset of products) becomes exceedingly complex in the presence of a large number of goods.

A natural question is then whether simpler pricing schemes are optimal under suitable demand conditions. A simple, commonly used mechanism consists of *upgrade pricing*, whereby the available options are ranked by set inclusion, i.e., some goods are only available as add-ons, Ellison (2005). For example, many online streaming services use a tiered subscription model, whereby users can pay to upgrade to a “premium package”—a subscription with a larger selection of the provider’s content relative to the “basic package”, Philips (2017).

In this paper, we obtain sufficient conditions under which upgrade pricing maximizes the seller’s revenue. Our approach consists of first identifying conditions under which the consumer’s types can be ordered in terms of their absolute or relative willingness to pay for the seller’s goods, and then ranking the goods themselves by the profitability of selling them to larger sets of consumer types. Our sufficient conditions not only establish the optimality of *some* upgrade pricing menu: they also

show that the optimal bundles are deterministic, and they reveal the order in which they are ranked in the menu. That is, we identify all the nested bundles that appear in the seller’s menu, and the profit-maximizing price for each one.

Our results consist of two distinct sets of conditions. The first set of conditions (Theorem 1) illustrates the essence upgrade pricing optimality in what we label as “regular” settings. While these conditions are reminiscent of regularity in one dimension, they are in fact weaker than the monotonicity of the buyer’s multidimensional types and of the (item by item) Myersonian virtual values. What we require is for the consumer’s types to be ranked in such a way that the virtual values for each item are negative over an initial and positive over a final segment. Furthermore, we require any consumer with a positive virtual value for an item to also have a larger value for that item, relative to any type with a negative virtual value. At the optimal prices, the lowest type buying each good is indifferent between buying it and not buying it. Finally, the sets of types buying each item are nested under the *weak* monotonicity property, which implies the optimal allocation can be implemented via upgrade pricing.

The second set of conditions (Theorem 2) describes our best attempt at extending our approach to non-regular distribution of types. In order to further weaken the regularity requirement, we restrict attention to type spaces for which the relative preference ratios for any two goods are monotone across types. An example of ordered relative preferences is if higher types have a stronger preference for good 2 over good 1. We refer to such a condition as “monotone marginal rates of substitution” (monotone MRS).

The intuition for our two results can be grasped by considering the demand functions for each good separately. Under monotonicity and monotone MRS, the optimal monopoly prices for each of the goods are ranked. In the special case where the Myersonian virtual values for our ordered types

$$\phi_i^k = \theta_i^k - \frac{1 - F_i}{f_i} (\theta_{i+1}^k - \theta_i^k)$$

are also monotone for each item  $k$ , the first set of conditions applies.

When virtual values are not monotone, however, they can cross zero more than once. In that case, the result still holds, but the proof requires the right ironing procedure. Our ironing procedure relaxes the standard approach of Myerson (1981) and the literature up to Haghpanah and Hartline (2020). Specifically, we do not iron with the goal of monotone virtual values, which corresponds to a concave revenue curve. Rather we iron towards single-crossing virtual values which leads to a *quasi*concave revenue curve. We then use the structure implied by monotone MRS to derive a dual certificate of optimality.

Under either set of conditions, each good is purchased by the same set of buyers that would buy it if that were the seller’s *only* product. We further show (Theorem 3) that, if the consumer’s types are (not weakly) monotone, the seller can equivalently post the vector of single-item monopoly prices—i.e., bundling is redundant. For example, in the case of two goods sold separately, monotone type spaces mean that no consumer type will buy good 2 without also buying good 1. More generally, the seller benefits from restricting the set of bundles the consumer can purchase through a proper menu of options with the upgrade property. However, examples also show that implementability through separate pricing is neither necessary nor sufficient for the optimality of upgrade pricing.

## 1.2 Related Literature

First and foremost, our paper contributes to the economics literature on product bundling. The profitability of mixed bundling relative to separate pricing was first examined by Adams and Yellen (1976), and further generalized by McAfee et al. (1989). More recently, a number of contributions

have studied the optimal selling mechanisms in the case of two or three goods, and derived conditions for the optimality of pure bundling (see, for example, Manelli and Vincent (2006) and Pavlov (2011)). Daskalakis et al. (2017) use duality methods to characterize the solution of the multiproduct monopolist’s problem, and show how the optimal mechanism may involve a continuum of lotteries over items. Bikhchandani and Mishra (2020) derive conditions under which the optimal mechanism is deterministic when the buyer’s utility is not necessarily additive. Finally, Ghili (2021) establishes conditions for the optimality of pure bundling when buyers’ values are interdependent. Relative to all these papers, we focus on a specific class of simple mechanisms, which includes pure bundling as a special case.

Hart and Nisan (2017) and Babaioff et al. (2014) also study the properties of simpler schemes. The former derives a lower bound on the revenue obtained from separate item pricing. The latter obtains an upper bound on the revenue of the optimal mechanism, relative to the better of pure bundling and separate pricing.

In the context of nonlinear pricing, Wilson (1993) suggested a "demand profile" approach that determines the price of each incremental unit by treating it as separate market. This approach is particularly attractive in settings where there is a natural ordering over the items. This in particular is the case when there is a homogeneous good that is offered in various quantities, such as in energy markets for electricity or water. This approach naturally generates a sequence of upgrade prices. The demand profile approach, and in particular the incremental pricing rule implied by it, does not always yield an optimal mechanism as consumers may wish to obtain earlier units in order to obtain the later units. Thus, a contribution of the current paper is to determine when upgrade pricing is exactly optimal and then to find the upgrade prices as solutions to the global revenue maximization problem rather than the incremental item problem. Other papers make assumptions that make sure that a demand profile-type approach yields an optimal mechanism. In Johnson and Myatt (2003), buyers have unit demand and sellers offer different varieties of a single good. The approach in their paper is to assume a quality ranking on the varieties and to solve for the upgrade prices—the additional payments required to buy a better variety. The survey of the nonlinear pricing literature by Armstrong (2016) covers related approaches that optimize upgrades separately.

Our formulation of the dual problem follows Cai et al. (2016), who present a general duality approach to Bayesian mechanism design. Cai et al. (2016) formulate virtual valuations in terms of dual variables, state the weak and the strong duality results, and use them to establish lower bounds for relative performance of simple mechanisms. An important contribution by Haghpanah and Hartline (2020) exploits the duality machinery to provide sufficient conditions for the exact optimality of a specific, simple mechanism—pure bundling—consisting of offering a maximal bundle at a posted price. Under their sufficient conditions, the dual variables can be recovered from a single-dimensional problem in which the seller is restricted to bundle all items together.

We follow the approach of Haghpanah and Hartline (2020) by leveraging the duality approach to provide sufficient conditions for the optimality of a particular class of mechanisms. Haghpanah and Hartline (2020) gave a characterization of the optimality of the grand bundle, we provide a characterization for upgrade pricing. As upgrade pricing allows multiple bundles to be present in the menu, we cannot assign the dual variables by solving a one-dimensional problem. Instead, we develop a novel ironing algorithm that generates these variables by ironing different item’s revenue curves for different types. Under our sufficient conditions, the so-constructed virtual surplus is maximized by an element-wise monotone allocation that can be implemented by upgrade pricing; by complementary slackness, this certifies the optimality of upgrade pricing. Because pure bundling is one instance of upgrade pricing, our conditions differ from those of Haghpanah and Hartline (2020).

Our ironing differs from existing ironing approaches using duality and tackles a more general problem. In comparison to Haghpanah and Hartline (2020), we prove optimality for mechanisms with menu size surpassing two. Fiat et al. (2016) studies a two-parameter model, and uses an ironing approach that leads from the revenue curves to their concave closure. Devanur et al. (2020) generalizes Fiat et al. (2016) to more general orders on the second parameter. Our approach tackles optimality for an arbitrary finite number of items and varies the ironing procedure. On a technical level, our ironing procedure yields quasi-concave ironed revenue curves, whereas the ironed revenue curves in Haghpanah and Hartline (2020); Fiat et al. (2016); Devanur et al. (2020) are concave.

Our results also feed into a literature specifying optimal finite mechanisms for multi-dimensional types. (Daskalakis et al., 2017, Section 7) for example characterizes the optimal mechanisms for the two-good monopolist problem if the optimal mechanism has a particular structure. While Daskalakis et al. (2017) requires that the region of the type space that is not allocated any item is not adjacent to all regions getting specific constant allocations, upgrade pricing mechanisms consistently break this requirement.

### 1.3 Structure of the Paper

The model is introduced in section 2. The first set of sufficient condition is presented in section 3. In section 4, we present our results for monotone MRS type spaces. In section 5, we discuss the relationship between separate pricing and upgrade pricing. We conclude in section 6.

## 2 Model

We consider a standard multiple-good monopoly setting. There is a single seller of  $d \geq 1$  goods and a single buyer. The seller's marginal costs of production are normalized to zero. The buyer's utility function is additive across goods. We refer to the vector of marginal utilities  $\theta_i \in \mathbb{R}^d$  as the buyer's type. Therefore, the utility of buyer type  $\theta_i$  from the consumption vector  $q \in [0, 1]^d$  is given by

$$U(\theta_i, q) = \sum_{k=1}^d \theta_i^k q^k.$$

We also adopt the shorthand notation  $\langle \theta_i, q \rangle := \sum_{k=1}^d \theta_i^k q^k$ . As a convention, we denote types by subscripts and items by superscripts. The buyer's utility is quasi-linear in transfers and her outside option is also normalized to zero.

The buyer knows her type. From the seller's perspective, the buyer's type is distributed over a finite set  $\Theta \subseteq \mathbb{R}_+^d$ , with  $|\Theta| = n$ , according to the distribution  $f \in \Delta(\Theta)$ . For any positive integer  $n$ , we adopt the convention that  $[n] := \{1, 2, \dots, n\}$ , and we index types by  $i \in [n]$ . We let  $f_i := f(\theta_i)$  and denote the cumulative distribution sequence by  $F_i = \sum_{j=1}^i f_j$ ,  $i \in [n]$ .

The seller aims to maximize revenue. By the revelation principle, we can focus on direct mechanisms  $(q, t) = (q_i, t_i)_{i \in \{0\} \cup [n]}$ . These mechanisms can be interpreted as menus with  $n + 1$  items so that item  $i$  delivers consumption vector  $q_i$  at price  $t_i$  and item  $(q_0, t_0) := (0, 0)$  captures the buyer's outside option.

We call a menu *upgrade pricing* if  $\{q_0, q_1, \dots, q_n\}$  can be ordered in the component-wise partial order on  $\mathbb{R}^d$  given by  $q \leq q' \Leftrightarrow \forall k \in [d]: q^k \leq (q')^k$ . Our main goal is to provide conditions under which upgrade pricing maximizes the seller's revenue among all direct mechanisms.

### 3 Optimal Mechanisms for Regular Distributions

We will make prominent use of the (partial) Lagrangian duality-based certificate of optimality used by Cai et al. (2016). We state the underlying duality result to fix notation.

#### 3.1 Duality

In what follows, we will associate with  $\lambda_{ji}$  the Lagrange multiplier of the incentive compatibility constraint of type  $\theta_j$  deviating to type  $\theta_i$ ,  $j \in [n], i \in \{0\} \cup [n]$ :

$$\langle q_j, \theta_j \rangle - t_j \geq \langle q_i, \theta_j \rangle - t_i.$$

We note that the incentive constraints corresponding to  $\lambda_{j0}$ ,  $j \in [n]$  are type  $j$ 's individual rationality constraints. As a main tool in our analysis, we define the multi-dimensional *virtual values* associated with Lagrange multipliers  $\lambda \in \mathbb{R}^n \times \mathbb{R}^{n+1}$  as

$$\phi_i^\lambda := \theta_i - \frac{1}{f_i} \sum_{j=1}^n \lambda_{ji} (\theta_j - \theta_i). \quad (1)$$

**Lemma 1.** *A mechanism  $(q_i, t_i)_{i \in \{0\} \cup [n]}$  maximizes revenue if and only if there exist multipliers  $\lambda_{ji}$ ,  $j \in [n], i \in \{0\} \cup [n]$  such that*

1.  $\lambda_{ji} \geq 0$  (*Non-Negativity*)
2.  $(q_i)_{i \in [n]}$  optimizes  $\max_{(q_i)_{i \in [n]} \in [0,1]^n} \sum_{i=1}^n f_i \langle q_i, \phi_i^\lambda \rangle$  (*Virtual Welfare Maximization*)
3.  $f_i = \sum_{j=0}^n \lambda_{ij} - \sum_{j=1}^n \lambda_{ji}$  for all  $i \in [n]$  (*Feasibility of Flow*)
4.  $\lambda_{ji} (\langle q_j, \theta_j \rangle - t_j - \langle q_i, \theta_j \rangle - t_i) = 0$  for all  $j \in [n], i \in \{0\} \cup [n]$  (*Complementary Slackness*)
5. *There are transfers  $t$  such that  $(q, t)$  is incentive compatible and individually rational (Implementability)*

We call the dual variables  $\lambda_{ji}$ ,  $j \in [n], i \in [n] \cup \{0\}$  *flows* from type  $j$  to type  $i$  whenever they are non-negative and satisfy Lemma 1 item 3. This name is inspired by flow conservation constraints from the maximum flow and minimum cost flow problem in discrete mathematics (Korte and Vygen, 2011).

*Proof of Lemma 1.* Slater's condition for affine inequality constraints (Boyd and Vandenberghe, 2004, p. 227) allows us to write revenue maximization subject to the incentive compatibility and individual rationality constraints as an unconstrained optimization problem for  $(q_i, t_i)_{i \in [n]}$  subject to complementary slackness and non-negativity of dual variables. The Lagrangian reads:

$$\begin{aligned} \mathcal{L} &= \sum_{i=1}^n f_i t_i + \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} (\langle q_j, \theta_j \rangle - t_j - \langle q_i, \theta_j \rangle - t_i) \\ &= \sum_{i=1}^n t_i \left( f_i - \sum_{j=0}^n \lambda_{ij} + \sum_{j=1}^n \lambda_{ji} \right) + \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_j, \theta_j \rangle - \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_i, \theta_j \rangle \\ &= \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_j, \theta_j \rangle - \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_i, \theta_j \rangle \\ &= \sum_{j=1}^n \left( \left( \sum_{i=1}^n \lambda_{ij} - \sum_{i=0}^n \lambda_{ji} \right) \langle q_j, \theta_j \rangle - \sum_{i=0}^n \lambda_{ji} (\langle q_i, \theta_j \rangle - \langle q_j, \theta_j \rangle) \right) \end{aligned}$$

$$= \sum_{j=1}^n \left( f_j \langle q_j, \theta_j \rangle - \sum_{i=0}^n \lambda_{ji} (\langle q_i, \theta_j \rangle - \langle q_j, \theta_j \rangle) \right) = \sum_{j=1}^n f_j \langle q_j, \phi_j \rangle.$$

Clearly, it is necessary for an optimal mechanism to be implementable. To conclude the proof, we need to show that virtual welfare maximization and feasibility of flow are equivalent to maximizing the Lagrangian. Assume virtual welfare maximization and feasibility of flow. Then the above equalities show that the Lagrangian is maximized and certify the optimality of the mechanism  $(q_i, t_i)_{i \in [n]}$ . Conversely, assume that the Lagrangian is maximized by  $(q_i, t_i)_{i \in [n]}$ . If the flow  $\lambda_{ji}$  were not feasible, then choosing  $t_i$  arbitrarily large or small would lead to a higher value for the Lagrangian, which yields a contradiction. Given that this is zero, the Lagrangian equals virtual welfare, and virtual welfare maximization follows from optimality.  $\square$

### 3.2 A Sufficient Condition for Regular Distributions

Our first set of sufficient conditions for upgrade pricing optimality consists of a weak monotonicity condition and a regularity condition.

We call a type distribution  $F$  *weakly monotone* with respect to cutoffs  $i^1, i^2, \dots, i^d \in [n]$  if for any  $i, j \in [n]$  and  $k \in \{1, 2, \dots, d\}$ ,

$$i \leq i^k \leq j \implies \theta_i^k \leq \theta_j^k.$$

Note that weak monotonicity is strictly weaker than monotonicity: for each item, only order comparisons with respect to a cutoff type need to hold, whereas types above or below the cutoff can be arbitrarily ordered.

Similarly, a type distribution  $F$  is *regular* with respect to cutoffs  $i^1, i^2, \dots, i^d \in [n]$  if for any  $i, j \in [n]$  and  $k \in \{1, 2, \dots, d\}$ ,

$$i \leq i^k \leq j \implies \phi_i^k \leq 0 \leq \phi_j^k, \quad (2)$$

where  $\phi_i$  denotes the *initial  $d$ -dimensional virtual values*

$$\phi_i := \theta_i - \frac{1 - F_i}{f_i} (\theta_{i+1} - \theta_i). \quad (3)$$

The initial  $d$ -dimensional virtual values can be seen as multi-dimensional versions of the virtual values in Myerson (1981).

We say that a type distribution  $F$  is *compatibly* weakly monotone and regular if it is both weakly monotone and regular with respect to the same set of cutoffs. When such cutoffs  $i^k$  exist, they are essentially unique except between contiguous types of vanishing virtual value  $\phi_i^k$  and monotone types  $\theta_i^k$ ,  $i \in [n]$ ,  $k \in [d]$ . Subfigure 1a illustrates a type distribution with this property.

Our regularity condition can be equivalently stated in terms of the *pseudo-revenues*

$$R_i^k := (1 - F_{i-1})\theta_i^k. \quad (4)$$

Subfigure 1c depicts pseudo-revenues. We call (4) *pseudo-revenue* because, without an assumption that the values are monotone with respect to the component-wise partial order, the pseudo-revenue does not correspond to the revenue from sales of item  $k$  at a posted price of  $\theta_i^k$ . In particular, because we have

$$\frac{R_i^k - R_{i+1}^k}{f_i} = \frac{(1 - F_{i-1})\theta_i^k - (1 - F_i)\theta_{i+1}^k}{f_i}$$

$$= \frac{f_i \theta_i^k - (1 - F_i)(\theta_{i+1}^k - \theta_i^k)}{f_i} = \theta_i^k - \frac{1 - F_i}{f_i}(\theta_{i+1}^k - \theta_i^k) = \phi_i^k, \quad (5)$$

imposing regularity with respect to the cutoffs  $i^k$  is equivalent to requiring that  $R_i^k$  is single-peaked with peak  $i^k$ . While pseudo-revenues do not have immediate economic meaning, they are an important technical tool, in particular for our analysis of non-regular distributions in section 4.

**Theorem 1.** *If the type distribution  $F$  is compatibly weakly monotone and regular with respect to cutoffs  $(i^k)_{k \in [d]}$ , then upgrade pricing is optimal. In particular, the following mechanism is optimal:*

$$q_i^k := \begin{cases} 1 & i \geq i^k \\ 0 & \text{else.} \end{cases}, \quad i \in [n], k \in [d]. \quad (6)$$

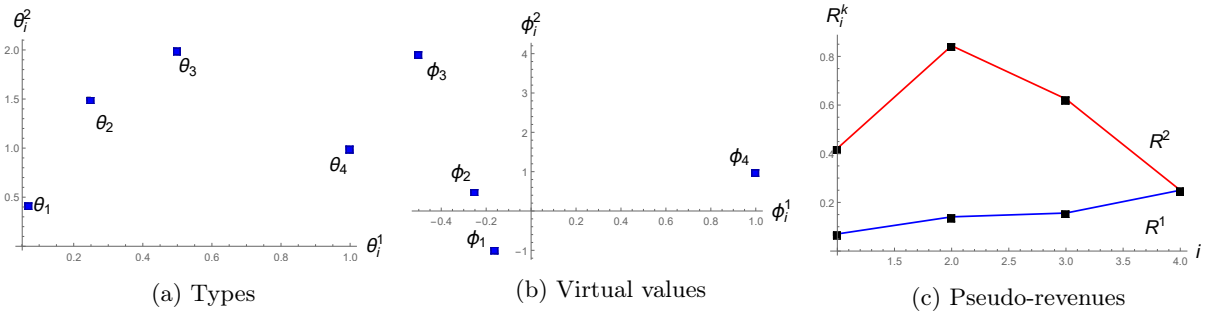


Fig. 1: Types, virtual values and pseudo-revenues for type space  $\Theta = \{(9/128, 27/64), (1/4, 3/2), (1/2, 2), (1, 1)\}$  and type distribution  $f = (7/16, 3/16, 1/8, 1/4)$ . The optimal mechanism sells good 2 at a price of 1 and good 1 as an upgrade, also at a price of 1. All types except  $\theta_1$  buy good 2, and only type  $\theta_4$  buys good 1.

*Proof of Theorem 1.* Define the dual variables

$$\hat{\lambda}_{ji} = \begin{cases} 1 - F_i & \text{if } j = i + 1 \\ 0 & \text{else.} \end{cases} \quad (7)$$

Observe that, by definition,  $\hat{\lambda}$  induces the initial virtual values,  $\phi_i = \phi_i^{\hat{\lambda}}$ .

We check the properties of Lemma 1. Virtual welfare maximization, condition 2, follows from

$$q_i^k = 1 \stackrel{(6)}{\iff} R_i^k \geq R_{i+1}^k \stackrel{(5)}{\iff} \phi_i^k \geq 0.$$

For flow preservation, condition 3, observe that

$$\sum_{j=1}^n \hat{\lambda}_{ij} - \sum_{j=0}^n \hat{\lambda}_{ji} = 1 - F_{i-1} - (1 - F_i) = f_i.$$

The mechanism is implementable, condition 5, by assumption of compatible weak monotonicity and regularity.

Finally, we need to check that complementary slackness (condition 4) holds. Observe that  $\hat{\lambda}_{ij} > 0$  implies  $j = i - 1$ . Hence, all types must be indifferent between their allocation and payment and

the allocation and payment of the next lower type. If the next lower type has the same allocation and payment, this is clearly satisfied. Otherwise, this is the first type buying an upgrade. If this type were not indifferent between buying it and not buying it, the price of the upgrade could be raised, and the revenue increased, without affecting other types' incentives. Thus, this type must be indifferent between their allocation (and payment) and the next lower type's allocation.  $\square$

Our assumptions of regularity and weak monotonicity relax the monotonicity of types and Myersonian virtual values by allowing for permutations above and below the monopoly price. These assumptions nonetheless require that the set of types that buy each object remains an upper selection, and conversely the set of types that do not buy remains a lower selection. The intuition for why this works is similar to the idea that the monopoly price does not depend on the valuations of types that are not marginally buying, just as long as they do not become marginal buyers.

These assumptions depend on the fixed order of types we have introduced in the model. Thus, if there exists an order that satisfies these assumptions, upgrade pricing is optimal. Furthermore, multiple orders of types might satisfy the theorem's conditions for a given type distribution  $F$ . In this case, the theorem can be used to certify optimality of mechanism (6), based on the different orders. As optimality of a mechanism for a distribution  $F$  does not depend on the order on types, the revenue of (6) must be the same for all orders with which the conditions of Theorem 1 are satisfied.

Our next set of conditions imposes similar requirements, strengthened appropriately to allow for non-regular type distributions, which require ironing.

#### 4 Optimal Mechanisms for Non-Regular Distributions

We now establish the optimality of an upgrade pricing mechanism in settings without regularity. The weaker sufficient conditions will replace the regularity condition and will allow for ironing to be part of the optimal mechanism. The new sufficient conditions will serve to allow us to perform the ironing procedure item-by-item, and limit the interaction of constraints across items. We say that a type space  $\Theta$  has *monotone marginal rates of substitution* if

$$1 \leq i \leq j \leq n \text{ and } 1 \leq k \leq l \leq d \implies \frac{\theta_i^l}{\theta_i^k} \leq \frac{\theta_j^l}{\theta_j^k}.$$

for any  $i, j \in [n]$ ,  $l, k \in [d]$ .

Recall that *pseudo-revenue* is given by  $R_i^k = (1 - F_i)\theta_i^k$ . We call a scalar sequence  $(R_i)_{i \in [n]}$ , *quasi-concave* if there is a cutoff  $i' \in [n]$  such that  $i' \leq i \leq j$  or  $j \leq i \leq i'$  implies  $R_i \geq R_j$ . We call the point-wise smallest quasi-concave sequence that point-wise dominates  $(R_i)_{i \in [n]}$  its *quasi-concave closure* and denote it by  $(\bar{R}_i)_{i \in [n]}$ . We will make regular use of the sequence  $(\bar{R}_i^k)_{i \in [n]}$ , the quasi-concave closure of the pseudo-revenue for item  $k$ .

To allow for our construction of a dual certificate of optimality, we need additional assumptions. These will be formulated in terms of *candidate ironing intervals*. For a pseudo-revenue  $R$ , we call a set of contiguous types  $I \subseteq [n]$  with

$$\bar{R}_i^k \neq R_i^k \tag{8}$$

for all  $i \in I$  such that there is no superset of contiguous types  $I' \supseteq I$  such that (8) holds for all  $i \in I'$ , a *candidate ironing interval for item  $k$* . (With slight abuse of language, we refer to discrete sets of contiguous types as *intervals*.) Every item  $k$  may have several candidate ironing intervals, and every type can be contained in a candidate ironing interval for different items.



We relax the regularity assumption on pseudo-revenues  $R_i^k$ . Instead of assuming regularity, i.e.  $R_i^k$  to be single-peaked with peak  $i^k$ , we assume two properties that are in combination weaker than regularity. We call a type distribution  $F$  *mostly regular* if for some cutoffs  $i^k \in \arg \max_{i \in [n]} R_i^k$  and any  $i$  such that  $i^k < i \leq i^{k+1}$ , the following hold:

1. (No partial overlap) If  $I$  is a candidate ironing interval of item  $k$  and  $J$  is a candidate ironing interval of item  $k + 1$ , then either  $I \cap J = \emptyset$  and there is  $i \in [n]$  such that  $I < i < J$  or  $J < i < I$ , or one of  $I, J$  is a subset of the other excluding its endpoints.
2. (No ironing on neighboring maxima) For any ironing candidate interval  $I$  of item  $k$ ,  $i^k, i^{k+1} \notin I$ .
3. (Not too shuffled) For any candidate ironing interval  $I \subseteq \{i^k + 1, i^k + 2, \dots, i^{k+1} - 1\}$  and  $i \in I$ ,

$$\theta_{\min I}^{k+1} \leq \theta_i^{k+1} \qquad \theta_{\max I}^k \leq \theta_{\max I+1}^k$$

Finally, we call a distribution *compatibly* weakly monotone and mostly regular if it is weakly monotone and mostly regular with respect to the same cutoffs  $i^k$ ,  $k \in [d]$ .

Note that monotone MRS by itself is not a restrictive assumption. For example, in two dimensions, every type set can be ordered in order of monotone MRS. In combination with compatible weak monotonicity and mostly regularity, this assumption becomes stronger.

Subfigure 2a shows the type space of a compatibly weakly monotone and mostly regular type distribution, and subfigure 2b its pseudo-revenues. Conversely, Figure 3 shows an instance of a

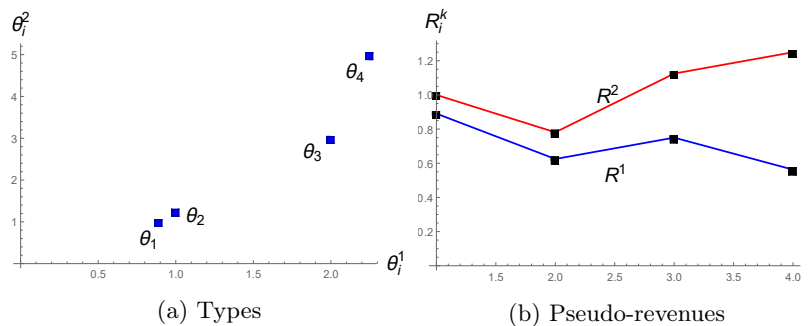


Fig. 2: Type space and pseudo-revenues for type space  $\Theta = \{(57/64, 1), (1, 5/4), (2, 3), (9/4, 5)\}$  and type distribution  $f = (3/8, 1/4, 1/8, 1/4)$ . The optimal mechanism sells good 1 at a price of  $57/64$ , and good 2 as an upgrade at a price of 5. All types buy good 1, and only type  $\theta_4$  buys good 2.

distribution over a monotone MRS type space that is *not* mostly regular. In particular, this example fails the first condition, because it involves overlapping candidate ironing intervals. We will use our assumptions to construct dual variables  $(\lambda_{ij})_{i,j \in [n]}$  by ironing pseudo-revenues for each item. In our proof that there is an optimal mechanism with an upgrade pricing allocation, we will use monotone MRS to show that for each type, ironing is only needed for two items, the lowest item in the MRS order that the type bought, and the highest item in the MRS order that she didn't buy. We will use the first two conditions of mostly regularity to show that from these two items, we can select a single item to iron at a time, while not changing the other item's virtual values in a way that will break virtual welfare maximization of the allocation. As in Theorem 1, weak monotonicity ensures implementability of an upgrade pricing allocation, i.e., the existence of a price vector  $(t_i)_{i \in [n]}$  such that the mechanism  $(q, t)$  is incentive compatible and individually rational. To allow for our ironing procedure to work, we also need a mild requirement on the monotonicity of types beyond weak monotonicity. While weak monotonicity was a requirement that could be formulated item-by-item, this requirement links the type order of neighboring items.

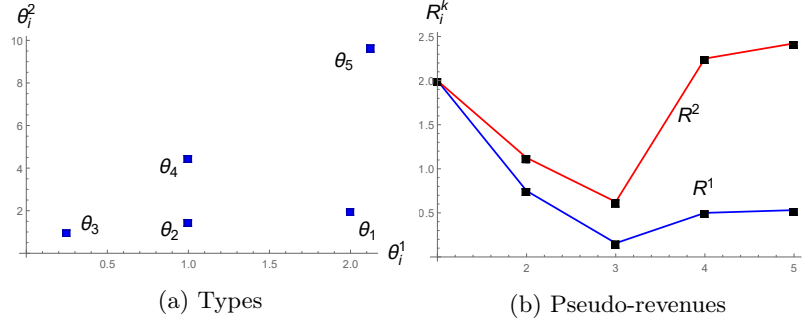


Fig. 3: Failure of no overlap:  $\{2, 3\}$  is a candidate ironing interval for item 2,  $\{3, 4\}$  is a candidate ironing interval for item 1.

**Theorem 2.** *Let  $\Theta$  have monotone marginal rates of substitution. If the type distribution  $F$  is compatibly weakly monotone and mostly regular with respect to cutoffs  $(i^k)_{k \in [d]}$ , then upgrade pricing is optimal. In particular, the following mechanism is optimal:*

$$q_i^k := \begin{cases} 1 & i \geq i^k \\ 0 & \text{else,} \end{cases} \quad i \in [n], k \in [d]. \quad (9)$$

Note that the allocation (9) is the allocation that arises from separate monopoly pricing.<sup>5</sup>

To prove Theorem 2, we will construct a sequence of flows  $\lambda_i$  from  $i = n$  down to  $i = 1$ , starting with  $\hat{\lambda}$ , the initial flow that induces Myersonian multi-dimensional virtual values. Given a definition of pseudo-revenue implied by a flow, our Ironing Algorithm will, for each type  $i$  and for at least one item  $k$ , iron to match induced pseudo-revenue with the quasi-concave closure of multi-dimensional Myersonian pseudo-revenue,  $\bar{R}_i^k$ . This is illustrated in Figure 4.

The main steps in this proof are to show that the ironing is well-defined in that such implied pseudo-revenue is attainable with a non-negative and feasible flow (Lemma 6 and Lemma 4, respectively). The most technical part of the proof consists of showing that the Ironing Algorithm produces dual variables that maximize virtual welfare (Lemma 3 and Lemma 7 (a)), and satisfy complementary slackness (Lemma 7 (b)).

Our first lemma is a main structural tool to link different items' virtual values and is tightly connected to monotone MRS. For  $k \in [d]$ ,  $i \in [n]$ , and flow  $\lambda$ , denote the normalized virtual value by

$$\nu_i^{\lambda, k} := \frac{\phi_i^{k, \lambda}}{\theta_i^k}.$$

The property that we will use repeatedly is that  $\nu_i^{\lambda, k}$  has the same sign as  $\phi_i^{\lambda, k}$ . We call a flow *downward* if  $\lambda_{ji} > 0$  for  $i, j \in [n]$  implies that  $j > i$ .

**Lemma 2.** *Let  $\Theta$  have monotone MRS. For any non-negative downward flow  $\lambda$ ,  $\nu_i^{\lambda, k} \geq \nu_i^{\lambda, l}$  for any  $1 \leq k \leq l \leq d$  and  $i \in [n]$ .*

*Proof.* It follows from definitions and monotone marginal rates of substitution that

$$\frac{\phi_i^{\lambda, k}}{\theta_i^k} = \frac{\theta_i^k - \sum_{j=1}^n \lambda_{ji} (\theta_j^k - \theta_i^k)}{\theta_i^k} = 1 + \sum_{j=i}^n \lambda_{ji} - \sum_{j=i}^n \lambda_{ji} \frac{\theta_j^k}{\theta_i^k}$$

<sup>5</sup> In section 5, we further explore the relationship between upgrade pricing and separate pricing, by showing conditions under which the allocation (9) can be implemented by a vector of single-item prices.

$$\geq 1 + \sum_{j=i}^n \lambda_{ji} - \sum_{j=i}^n \lambda_{ji} \frac{\theta_j^l}{\theta_i^l} = \frac{\theta_i^l - \sum_{j=1}^n \lambda_{ji}(\theta_j^l - \theta_i^l)}{\theta_i^l} = \frac{\phi_i^{\lambda,l}}{\theta_i^l}.$$

□

The next Lemma shows that virtual welfare maximization reduces to virtual welfare maximization for the neighboring items, i.e., the last item that a type buys and the first item that a type does not buy—with respect to the MRS order.

**Lemma 3.** *Assume  $\Theta$  has monotone MRS and mostly regular and that there exists a non-negative downward flow  $\lambda$  such that for any  $i \in [n]$  such that  $i^k \leq i \leq i^{k+1}$ , we have  $\phi_i^{\lambda,k} \geq 0$  and  $\phi_i^{\lambda,k+1} \leq 0$ . Then, the allocation in (9) maximizes virtual welfare.*

*Proof.* Fix  $i \in [n]$  such that  $i^k \leq i \leq i^{k+1}$ . Note that as  $\phi_i^{\lambda,k}$  and  $\nu_i^{\lambda,k}$  are positive multiples of each other, Lemma 2 implies the implications

$$\begin{aligned} \phi_i^{\lambda,k+1} \leq 0 &\implies \phi_i^{\lambda,l} \leq 0, & l \geq k+1 \\ \phi_i^{\lambda,k} \geq 0 &\implies \phi_i^{\lambda,l} \geq 0, & l < k. \end{aligned}$$

Therefore, the assumption implies that  $\phi_i^{\lambda,l} \leq 0$  for any  $l > k$  and  $\phi_i^{\lambda,l} \geq 0$  for any  $l \leq k$ , which ensures virtual welfare maximization of (9). □

For  $k = 0$  and  $k = d$  this Lemma reduces virtual welfare maximization for all items, and ironing for all items, to virtual welfare maximization for the first resp. last item. Finding a flow that maximizes virtual welfare reduces to ironing the (one-dimensional) virtual values  $\phi_i^1$  and  $\phi_i^d$ . For types  $i \leq i^1$  and  $i \geq i^d$ , we can hence use techniques from one-dimensional ironing and iron the pseudo-revenue to its concave closure in a discrete variant of Myerson (1981)'s procedure. From now, our discussion therefore focuses on  $k \in [d-1]$  and  $i \in [i^k+1, i^{k+1}]$ , i.e. types where an ironing that ensures virtual welfare maximization for both item  $k$  and item  $k+1$  is needed.

The following algorithm will make use of  $\hat{\lambda}$  as defined in (7), the *initial flow* and of a generalization of the pseudo-revenue. The pseudo-revenue associated to a flow  $\lambda$ ,  $R_i^{\lambda,k}$  is

$$R_i^{\lambda,k} = \sum_{j=i}^n f_j \phi_j^{\lambda,k}.$$

This generalization is intuitive, as virtual values are, as in (5), slopes of pseudo-revenues

$$\frac{R_i^{\lambda,k} - R_{i+1}^{\lambda,k}}{f_i} = \frac{\sum_{j=i}^n f_j \phi_j^{\lambda,k} - \sum_{j=i+1}^n f_j \phi_j^{\lambda,k}}{f_i} = \phi_i^{\lambda,k}. \quad (10)$$

Our algorithm will adjust a flow by raising one point in a revenue sequence at a time, from right to left. We will prove that this will yield slopes of revenue sequences—i.e. virtual values—which have the correct sign for virtual welfare maximization of (9). This is non-trivial, as pseudo-

revenues for different items might not move in the same direction when dual variables are changed.

```

 $\lambda \leftarrow \hat{\lambda};$ 
for  $i = n$  to 1 do
  Let  $\gamma_i \in [0, 1]$  be maximal such that for
    
$$\begin{aligned} \lambda'_{ji} &\leftarrow \gamma_i \lambda_{ji}, & \forall j : n > j > i \\ \lambda'_{j(i-1)} &\leftarrow \lambda_{j(i-1)} + (1 - \gamma_i) \lambda_{ji}, & \forall j : n > j > i \\ \lambda'_{i(i-1)} &\leftarrow \lambda_{i(i-1)} - (1 - \gamma_i) \sum_{i'=i}^n \lambda_{i'i}, \end{aligned} \tag{11}$$

     $R_i^{\lambda', \kappa(i)} = \overline{R}_i^{\kappa(i)}$  holds;
     $\lambda \leftarrow \lambda';$ 
  Return  $\lambda'$ ;

```

**Algorithm:** Ironing, parameterized by an ironing mapping  $\kappa: [n] \rightarrow [d]$

The flow (11) was used earlier in Haghpanah and Hartline (2020). An important difference is that Haghpanah and Hartline (2020) choose  $\gamma_i$  to iron the revenue sequence of the grand bundle to the concave closure of pseudo-revenue. Instead, we iron to the *quasi*-concave closure of (their equivalent of) pseudo-revenue of an item  $\kappa(i)$ . The parameter  $\gamma_i$  can be found as solution to a system of linear equations. We show that a solution  $\gamma_i \in [0, 1]$  exists in Lemma 6.

We first observe that the Ironing Algorithm outputs a flow which is non-negative and feasible.

**Lemma 4.** *The output of the Ironing Algorithm is a flow, i.e. non-negative and satisfies flow preservation, Lemma 1 item 3.*

*Proof.* We prove the claim by induction from  $i = n$  to  $i = 1$ . Note that  $\hat{\lambda}$  is feasible as argued in the proof of Theorem 1, which starts the induction. Fix an arbitrary iteration  $i \in [n]$ , and assume that  $\lambda$  is feasible. We check that the difference in excess flow, i.e. incoming and outgoing flow, cancel out for  $j > i$ . For  $i$  and  $i + 1$  similar calculations yield the result. We omit these. For any  $j > i$ ,

$$\lambda'_{ji} - \lambda_{ji} + \lambda'_{j(i-1)} - \lambda_{j(i-1)} = \gamma \lambda_{ji} - \lambda_{ji} + \lambda_{j(i-1)} + (1 - \gamma_i) \lambda_{ji} - \lambda_{j(i-1)} = 0.$$

Now consider non-negativity. Each  $\lambda_{ij}$  reduces at most once during the course of the Ironing Algorithm. More specifically, only if  $j = i - 1$  and during iteration  $i$ . In this iteration,

$$\lambda'_{i(i-1)} = \lambda_{i(i-1)} - (1 - \gamma_i) \sum_{i'=i}^n \lambda_{i'i} \geq \lambda_{i(i-1)} - \sum_{i'=i}^n \lambda_{i'i} = f_i \geq 0,$$

where we used that  $\gamma_i \leq 1$ ,  $\lambda_{ir} = \hat{\lambda}_{ir}$ ,  $r < i - 1$ , which in particular implies that  $\lambda_{ir} = 0$ , and feasibility of the flow.  $\square$

Next observe that in the Ironing Algorithm, iteration  $i$  changes the revenue (for any item  $k$ ) only for type  $i$ . Hence, our ironing algorithm raises pseudo-revenue for one type at a time.

**Lemma 5.** *For any iteration  $i$ ,  $R_j^{\lambda', k} = R_j^{\lambda, k}$  for any  $j \neq i$ . In particular,  $\phi_j^{\lambda', k} = \phi_j^{\lambda, k}$  for  $j \notin \{i - 1, i\}$ .*

*Proof.* First note that as the in-flow for higher types remains unchanged in iteration  $i$ ,  $\lambda'_{rj} = \lambda_{rj}$ ,  $r \in [n]$ ,  $j > i$ , the revenue does not change,  $R_j^{\lambda', k} = R_j^{\lambda, k}$ . For types  $j < i$ , we check that the changes

to virtual welfare on the types whose inflows do change,  $i$  and  $i - 1$ , cancel out. By definition of virtual values,

$$\begin{aligned}\phi_i^{\lambda',k} &= \phi_i^{\lambda,k} + \frac{1-\gamma_i}{f_i} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k) \\ \phi_{i-1}^{\lambda',k} &= \phi_{i-1}^{\lambda,k} + \frac{1-\gamma_i}{f_{i-1}} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k) - \frac{1-\gamma_i}{f_{i-1}} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k) \\ &= \phi_i^{\lambda,k} - \frac{1-\gamma_i}{f_i} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k).\end{aligned}$$

Hence,

$$f_{i-1}\phi_{i-1}^{\lambda',k} + f_i\phi_i^{\lambda',k} = f_{i-1}\phi_{i-1}^{\lambda,k} + f_i\phi_i^{\lambda,k}.$$

The statement on the virtual values follows from (10).  $\square$

Before showing that  $\gamma_i$  in the algorithm always exists, we define the ironing function  $\kappa(i)$ .

By no ironing on neighboring maxima, each candidate ironing interval  $I$  must be contained in an interval  $\{i^k, i^k + 1, \dots, i^{k+1}\}$ . By this condition, in addition to no partial overlap, for each type  $i$ , there is a unique inclusion maximal candidate among the candidate ironing intervals for items  $k$  and  $k + 1$ . We let  $\kappa(i)$  denote the item this interval is a candidate ironing interval for. If  $i$  is not part of any ironing interval, we set  $\kappa(i)$  arbitrarily in  $\{k, k + 1\}$ . We call  $\kappa(i)$  the *ironed item for type  $i$*  and piece-wise constant intervals of  $\kappa$  *ironing intervals*.

**Lemma 6.** *Assume that  $F$  is mostly regular. Then, for each  $i \in [n]$ ,  $\gamma_i$  such that  $R_i^{\lambda_i(\gamma_i), \kappa(i)} = \bar{R}_i^{\kappa(i)}$  exists. In particular, the Ironing Algorithm is well-defined.*

*Proof.* We prove this statement by induction from  $i = n$  down to 1. Let  $i \in [n]$  and assume that  $R_{i+1}^{\lambda, \kappa(i)} = \bar{R}_{i+1}^{\lambda, \kappa(i)}$ . If  $i$  is not part of an ironing interval, then by definition of ironing intervals and Lemma 5,  $R_i^{\lambda, k} = \bar{R}_i^{\lambda, k}$ , and the induction step is trivial by choosing  $\gamma_i = 1$ , yielding  $R_i^{\lambda'(1), k} = \bar{R}_i^{\lambda'(1), k}$ . Otherwise,  $i$  is in an ironing interval. Let  $\kappa(i) = k$ . By no partial overlap, if  $i + 1$  is part of an ironing interval, it must be part of the same ironing interval, in particular must have been ironed for item  $k$ . Hence, by the induction hypothesis,  $R_{i+1}^{\lambda, k} = \bar{R}_{i+1}^{\lambda, k}$ .

Denote

$$\bar{\phi}_i^k = \frac{\bar{R}_i^k - \bar{R}_{i+1}^k}{f_i}$$

the slope of the quasi-concave closure of pseudo-revenue of item  $k$  at type  $i$ . By definition of the quasi-concave closure, the slope of the revenue curve must be non-positive,

$$\bar{\phi}_i^k \leq 0.$$

As all types are non-negative, we get that

$$\bar{\phi}_i^k \leq 0 \leq \theta_i^k = \phi_i^{\lambda'(0), k}. \quad (12)$$

Again by Lemma 5,  $\bar{R}_{i+1}^k = R_{i+1}^{\lambda'(0), k}$ . Therefore

$$\bar{R}_i^k = f_i \bar{\phi}_i^k + \bar{R}_{i+1}^k = f_i \phi_i^{\lambda'(0), k} + R_{i+1}^{\lambda'(0), k}$$

$$\leq f_i \phi_i^{\lambda'(0),k} + R_{i+1}^{\lambda'(0),k} = R_i^{\lambda_i(0),k}.$$

In particular,  $\overline{R}_i^k \leq R_i^{\lambda_i(0),k}$ .

Also, by Lemma 5 and the definition of the quasi-concave closure,  $R_i^{\lambda'(1),k} = R_i^{\lambda,k} \leq \overline{R}_i^k$ . As  $\gamma \mapsto R_i^{\lambda'(\gamma),k}$  is a continuous function, the existence of the desired  $\gamma \in [0, 1]$  follows from the Intermediate Value Theorem.  $\square$

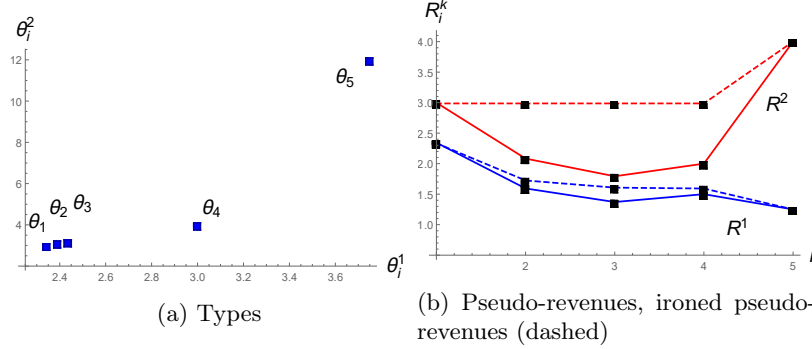


Fig. 4: Ironing of virtual values and corresponding pseudo-revenues.

The last lemma before the proof of Theorem 2 shows that the output of the algorithm satisfies complementary slackness and the condition of Lemma 3, which is sufficient for virtual welfare maximization.

**Lemma 7.** *Assume that  $\Theta$  is has monotone MRS, and that  $F$  is mostly regular. Then,  $q$  maximizes virtual welfare and satisfies the requirements of Lemma 3 with respect to  $\lambda'$ , the output of the Ironing Algorithm.*

*Proof.* We first show that  $\lambda'$  satisfies the requirements of Lemma 7, i.e. that for any  $i^k \leq i \leq i^{k+1}$  we have that  $\phi_i^{\lambda,k} \geq 0$  and  $\phi_i^{k+1,\lambda} \leq 0$ . Let  $i \in [n]$ . If  $i$  is not in an ironing interval, and there is no ironing interval  $I$  such that  $i = \min I - 1$ , the claim follows from the definition of the quasi-concave closure and our definition of ironing intervals, as well as Lemma 5.

We consider the remaining cases  $i = \max I$ ,  $i = \min I - 1$ , and  $i \in I \setminus \{\max I\}$  separately. For the third case, by definition of the ironing and the quasi-concave closure,

$$\phi_i^{\lambda',\kappa(i)} = \nu_i^{\lambda',\kappa(i)} = 0.$$

For  $\kappa(i) = k$ ,  $\nu_i^{k+1,\lambda} \leq 0$  by Lemma 2 and hence  $\phi_i^{k+1,\lambda} \leq 0$  as  $\nu_i^{k+1,\lambda}$  and  $\phi_i^{k+1,\lambda}$  are positive multiples of each other. Similarly, we have for  $\kappa(i) = k + 1$  that  $\nu_i^{k,\lambda} \geq 0$  by Lemma 2 and hence  $\phi_i^{k,\lambda} \geq 0$ . It remains to consider  $i = \min I - 1$  and  $i = \max I$ . We consider these separately for  $\kappa(i) = k$  and  $\kappa(i) = k + 1$ , for a total of four cases.

Case 1.  $i = \max I$  and  $\kappa(i) = k$ . By definition of the quasi-concave closure,  $\overline{R}_i^k = R_i^k$ , the algorithm chooses  $\gamma_i = 1$  and hence  $\phi_i^{\lambda',k} = \phi_i^k = 0$  and hence by arguments as above,  $\phi_i^{\lambda',k+1} = \phi_i^{k+1} \leq 0$ .

Case 2.  $i = \min I - 1$  and  $\kappa(i) = k + 1$ . By definition of the quasi-concave closure,  $\overline{R}_i^{k+1} = R_i^{k+1}$ , the algorithm chooses  $\gamma_i = 1$  and hence, by an argument similar to case 1,  $\phi_i^{\lambda,k} = \phi_i^k \geq 0$ .

Case 3.  $i = \max I$  and  $\kappa(i) = k + 1$ .

The derivative of the virtual value of the right end of the ironing interval, type  $i$  is given by

$$\frac{\partial \phi_i^{\lambda', k}}{\partial \gamma} = -\frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji}(\theta_j^k - \theta_i^k) = -\frac{1}{f_i} \lambda_{(i+1)i}(\theta_{i+1}^k - \theta_i^k). \quad (13)$$

By no partial overlap,  $0 \leq \phi_i^k$ . Also, by not too shuffledness, (13) is non-positive. Hence,  $\phi_i^k = \phi_i^{\gamma'(0), k} \leq \phi_i^{\lambda', k}$ , and the algorithm chooses  $\gamma_i < 1$ . Combining these observations, we obtain  $\phi_i^{k, \lambda'} \geq 0$ . Because the algorithm chooses  $\gamma_i < 1$ , this implies

$$\phi_i^{\lambda', k} \geq \phi_i^k \geq 0.$$

Case 4.  $i = \min I - 1$  and  $\kappa(i) = k$ . The derivative of the virtual value of the next item at the left end of the ironing interval, type  $i$  is given by

$$\begin{aligned} \frac{\partial \phi_i^{\lambda', k+1}}{\partial \gamma} &= \frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji}(\theta_j^{k+1} - \theta_i^{k+1}) - \frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji}(\theta_{i+1}^{k+1} - \theta_i^{k+1}) \\ &= \frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji}(\theta_j^{k+1} - \theta_{i+1}^{k+1}) \geq 0, \end{aligned} \quad (14)$$

where the last inequality follows from not-too-shuffledness (note that  $i + 1 = \min I$ ). By no partial overlap,  $0 \geq \phi_i^{k+1}$ . Moreover,  $\phi_i^{k+1} = \phi_i^{\gamma'(0), k} \geq \phi_i^{\lambda', k+1}$  because of (14) and because the algorithm chooses  $\gamma_i < 1$ . Combining these observations, we obtain  $\phi_i^{k+1, \lambda} \leq 0$ .

To show complementary slackness, observe that whenever  $i$  is not in an ironing interval, as  $\gamma_i$  is chosen maximal such that  $\bar{R}_i^{\kappa(i)} = R_i^{\kappa(i)}$ , the Algorithm chooses  $\gamma_i = 1$ , which implies that for  $j > i > r$  and for  $j > i + 1$  and  $i = r$ ,  $\lambda'_{jr} = 0$ . By no ironing over maxima, this implies that for  $j > i^k > r$ ,  $\lambda_{jr} = 0$ . Moreover, as an invariant of the algorithm the flow  $\lambda$  is downward. Hence, the only dual variables that are tight are within types that get the same allocation and payment (corresponding to  $\lambda_{ij}$  such that  $i^k \leq i, j \leq i^{k+1}$ ,  $k \in [d]$ ) or local downward constraints (corresponding to  $\lambda_{(i+1)i}$ ,  $i \in [n - 1]$ ). The former incentive constraints clearly bind, the latter bind by weak monotonicity, as for the marginally buying type (which, by compatibility of weak monotonicity with mostly regularity must also be the first type in the MRS order), the price could be raised if she were not indifferent between her allocation and payment and the allocation and payment of the next lower type.  $\square$

We are now ready to finish the proof of Theorem 2.

*Proof of Theorem 2.* Implementability follows from weak monotonicity and the definition of the optimal mechanism, (9). Non-negativity and feasibility of flow are properties of the Ironing Algorithm shown in Lemma 4. Virtual welfare maximization and complementary slackness have been shown in Lemma 7.  $\square$

## 5 Upgrade Pricing and Separate Pricing

In both Theorem 1 and Theorem 2, we established the optimality of an upgrade pricing mechanism that yields the same allocation as separate (item by item) monopoly pricing, though not necessarily the same transfers. We will show in this section that, under monotonicity with respect to the

component-wise partial order, separate pricing and upgrades become equivalent—upgrade pricing is redundant.

We say that the type space  $\Theta$  is *monotone* if  $\theta_i^k \leq \theta_j^k$  for any  $i < j \in [n]$  and  $k \in [d]$ .

We call a mechanism separate pricing if a type separately chooses whether to buy each item  $k$  at a price  $p_k$ . Formally, a mechanism satisfies separate pricing if it can be written as:

$$q_i^k = \begin{cases} 1 & \theta_i^k \geq p_k \\ 0 & \text{else,} \end{cases} \quad t_i = \sum_{k=1}^d p_k \mathbb{1}_{q_i^k=1}.$$

**Theorem 3.** *If the type space  $\Theta$  is monotone, then the outcome of any upgrade pricing mechanism can be implemented via separate pricing, and conversely. When the type space is not monotone, neither implication needs to hold.*

*Proof.* We first assume types are monotone and show that the allocation and revenue of any upgrade pricing mechanism can be obtained through a separate pricing mechanism, and vice versa.

Let  $\theta_1 \leq \dots \leq \theta_i \leq \dots \leq \theta_n$ , and fix an upgrade pricing mechanism  $\mathcal{M}$ . This mechanism admits an indirect representation as (a) a collection of bundles ranked by set inclusion  $\{b_k\}_{k=0}^K$ , with  $b_0 = \emptyset$  and  $K \leq d$ , and (b) a vector of prices  $t_k$  that are increasing in  $k$ , with  $t_0 = 0$ . Let  $\underline{\theta}_k$  and  $\bar{\theta}_k$  denote the lowest and highest types who choose bundle  $b_k$  under mechanism  $\mathcal{M}$ . Because types are monotone, buyer self-selection implies  $\underline{\theta}_k \geq \bar{\theta}_{k-1}$ .

We now construct a separate pricing mechanism, i.e., a vector of prices  $\{p_j\}_{j=1}^d$  that yields the same allocation and payments as our upgrade pricing mechanism. To do so, define the collection of *upgrades*  $u_k := b_k \setminus b_{k-1}$  and the upgrade prices  $\tau_k := t_k - t_{k-1}$ . For each upgrade bundle  $k$  and every good  $j \in u_k$ , let the single-item prices  $p_j$  satisfy

$$p_j \in \left[ \bar{\theta}_{k-1}^j, \underline{\theta}_k^j \right] \quad \text{and} \quad \sum_{j \in u_k} p_j = \tau_k.$$

Under monotonicity, such a vector of prices always exists. By consumer self-selection in the original mechanism  $\mathcal{M}$ , we have

$$\begin{aligned} \bar{\theta}_{k-1} b_k - t_k &\leq \bar{\theta}_{k-1} b_{k-1} - t_{k-1}, \\ \underline{\theta}_k b_{k-1} - t_{k-1} &\leq \underline{\theta}_k b_k - t_k. \end{aligned}$$

In turn, this implies

$$\bar{\theta}_{k-1} u_k \leq \tau_k \leq \underline{\theta}_k u_k.$$

With the prices so constructed, each type purchases the same goods as under  $\mathcal{M}$  and pays the same total price. Notice first that each type's choice from the original mechanism  $\mathcal{M}$  is still available at the same price, i.e., each bundle  $b_k$  can still be purchased for a total price  $t_k$ . Moreover, by monotonicity, no type  $\theta$  who buys bundle  $b_k$  under the upgrade pricing mechanism  $\mathcal{M}$  derives positive net surplus from any object  $j \in u_{k'}$  with  $k' > k$  under the separate prices constructed above. And finally, no such type  $\theta$  derives positive net surplus by removing any object  $j \in u_{k'}$  with  $k' \leq k$  from her consumption bundle.

The other direction of this result is immediate: if types are monotone, the goods purchased by two different types under any separate pricing mechanism are ranked by set inclusion. Thus, replacing the separate pricing mechanism with the resulting upgrade pricing mechanism yields the same outcome.



Finally, we show by means of two counterexamples that, without type monotonicity, separate pricing is not equivalent to upgrade pricing.

In particular, there exist type spaces and vectors of separate prices that do not induce an upgrade pricing allocation. For example, let  $\Theta = \{(1, 1), (1, 3), (3, 3), (4, 1)\}$  and consider the separate prices  $p = (2, 2)$ : type  $\theta_2$  buys good 2 only, type  $\theta_3$  buys both goods, and type  $\theta_4$  buys good 1 only.

Likewise, for the same type space, consider the upgrade pricing mechanism where  $q = (0, 1)$  is sold for  $t = 2$  and  $q = (1, 1)$  is sold for  $t = 4$ , i.e., good  $j = 1$  is only sold as an upgrade, for an additional price  $\tau = 2$ . Under this mechanism, type  $\theta_2$  buys good 2 only, while types  $\theta_3$  and type  $\theta_4$  buy both goods. However, as we saw above, the vector of separate prices  $p = (2, 2)$  yields a different allocation (and a lower revenue for the seller).  $\square$

Whenever an upgrade pricing mechanism implements the allocation of optimal separate pricing, each marginal type  $\underline{\theta}_k$  is indifferent by construction between the two consecutive bundles  $b_{k-1}$  and  $b_k$ . Theorem 3 then implies that the outcome of this mechanism can be implemented by the separate monopoly prices.

**Corollary 1.** *If  $\Theta$  is monotone,  $q$  is an allocation of an optimal upgrade pricing mechanism, and  $q$  is the allocation of separate monopoly pricing, then separate monopoly pricing is optimal.*

Adding a monotonicity condition to both of our main theorems, Theorem 1 and Theorem 2, we hence obtain two sets of sufficient conditions under which separate monopoly pricing is optimal.

**Corollary 2.** *If  $\Theta$  is monotone and  $F$  is regular, separate monopoly pricing is optimal.*

Note that monotone types are not too shuffled.

**Corollary 3.** *If  $\Theta$  is monotone and has a monotone marginal rates of substitution, and  $F$  is mostly regular, then separate monopoly pricing is optimal.*

## 6 Conclusion

It is a common practice for a seller to offer bundles of products or services that are ordered in a way that more expensive bundles contain all items from less expensive bundles as well as some extra items. In this paper, we provide sufficient conditions under which such “upgrade pricing” schemes are exactly optimal for a monopolist seller.

There are several ways in which the current analysis could be extended. First, our conditions could be relaxed to account for richer type spaces and type distributions, such as a continuum of types in the  $d$ -dimensional space. One natural extension can be obtained immediately: assume that a type distribution can be split into several type cohorts, in fact quantized type space, such that each type cohort satisfies the conditions of our theorems. Our results imply that the optimal mechanisms in each respective cohort are upgrade pricing. In this respect, Bergemann et al. (2021) show that in nonlinear pricing problems, the revenue of the continuous type space is generally well approximated by a finite quantized type space.

Second, our sufficient conditions for the optimality of upgrade pricing may be complemented by necessary conditions. In doing so, one may want to distinguish between conditions on type distributions and type spaces. For example, one may ask which type spaces guarantee that upgrade pricing is optimal irrespective of the type distribution.

Finally, throughout the paper we highlight the interplay between optimality of different pricing schemes: bundling, upgrade pricing, and separate sales. It would be instructive to provide a more complete characterization of the cases in which one of these schemes strictly outperforms another.

**Acknowledgements** We thank Mark Armstrong and the seminar audience at MIT for helpful comments. Bergemann and Bonatti acknowledge financial support through NSF SES 1948336. Smolin acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d’Avenir) program (grant ANR-17-EURE-0010).

## Bibliography

- W. J. Adams and J. L. Yellen. Commodity Bundling and the Burden of Monopoly. *Quarterly Journal of Economics*, 90(3):475–498, 1976.
- M. Armstrong. Nonlinear pricing. *Annual Review of Economics*, 8:583–614, 2016.
- M. Babaioff, N. Immorlica, B. Lucier, and S. M. Weinberg. A Simple and Approximately Optimal Mechanism for an Additive Buyer. *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 21–30, 2014. <https://doi.org/10.1109/focs.2014.11>.
- D. Bergemann, E. Yeh, and J. Zhang. Nonlinear pricing with finite information. *Games and Economic Behavior*, 130:62–84, 2021.
- S. Bikhchandani and D. Mishra. Selling two identical objects. *arXiv:2009.11545*, 2020.
- S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. ISBN 9780521833783. <https://doi.org/10.1198/jasa.2005.s41>.
- Y. Cai, N. R. Devanur, and S. M. Weinberg. A Duality Based Unified Approach to Bayesian Mechanism Design. *Proceedings of the Annual ACM Symposium on Theory of Computing*, 19–21 June, 2016. ISSN 07378017. <https://doi.org/10.1145/2897518.2897645>.
- C. Daskalakis, A. Deckelbaum, and C. Tzamos. Strong Duality for a Multiple-Good Monopolist. *Econometrica*, 85(3):735–767, 2017. ISSN 0012-9682. <https://doi.org/10.3982/ecta12618>.
- N. R. Devanur, K. Goldner, R. R. Saxena, A. Schwartzman, and S. M. Weinberg. Optimal mechanism design for single-minded agents. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 193–256, 2020.
- G. Ellison. A Model of Add-On Pricing. *Quarterly Journal of Economics*, 120(2):585–637, 2005.
- A. Fiat, K. Goldner, A. R. Karlin, and E. Koutsoupias. The fedex problem. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 21–22, 2016.
- S. Ghili. A Characterization for Optimal Bundling of Products with Interdependent Values. Technical report, Yale University, 2021.
- N. Haghpanah and J. Hartline. When Is Pure Bundling Optimal? *The Review of Economic Studies*, 88(3):1127–1156, 08 2020. ISSN 0034-6527. <https://doi.org/10.1093/restud/rdaa053>. URL <https://doi.org/10.1093/restud/rdaa053>.
- S. Hart and N. Nisan. Approximate revenue maximization with multiple items. *Journal of Economic Theory*, 172:313–347, 2017.
- J. P. Johnson and D. P. Myatt. Multiproduct quality competition: Fighting brands and product line pruning. *American Economic Review*, 93(3):748–774, 2003.
- B. H. Korte and J. Vygen. *Combinatorial optimization*, volume 1. Springer, 2011.
- A. M. Manelli and D. R. Vincent. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1–35, 2006.
- R. P. McAfee, J. McMillan, and M. D. Whinston. Multiproduct Monopoly, Commodity Bundling, and Correlation of Values. *Quarterly Journal of Economics*, 104(2):371–383, 1989.
- R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.
- G. Pavlov. Optimal mechanism for selling two goods. *The BE Journal of Theoretical Economics*, 11(1), 2011.
- J. Philips. Don’t look now, but the great unbundling has spun into reverse. *New York Times*, 2017.
- R. Wilson. *Nonlinear Pricing*. Oxford University Press, Oxford, 1993.