

# Simple Mechanisms for a Combinatorial Buyer and Applications to Revenue Monotonicity

Aviad Rubinstein\*      S. Matthew Weinberg†

May 7, 2019

## Abstract

We study the revenue maximization problem of a seller with  $n$  heterogeneous items for sale to a single buyer whose valuation function for sets of items is unknown and drawn from some distribution  $D$ . We show that if  $D$  is a distribution over subadditive valuations with independent items, then the better of pricing each item separately or pricing only the grand bundle achieves a constant-factor approximation to the revenue of the optimal mechanism. This includes buyers who are  $k$ -demand, additive up to a matroid constraint, or additive up to constraints of any downwards-closed set system (and whose values for the individual items are sampled independently), as well as buyers who are fractionally subadditive with item multipliers drawn independently. Our proof makes use of the core-tail decomposition framework developed in prior work showing similar results for the significantly simpler class of additive buyers [17, 1].

In the second part of the paper, we develop a connection between approximately optimal simple mechanisms and approximate revenue monotonicity with respect to buyers' valuations. Revenue non-monotonicity is the phenomenon that sometimes strictly *increasing* buyers' values for every set can strictly *decrease* the revenue of the optimal mechanism [14]. Using our main result, we derive a bound on how bad this degradation can be (and dub such a bound a proof of *approximate* revenue monotonicity); we further show that better bounds on approximate monotonicity imply a better analysis of our simple mechanisms.

---

\*UC Berkeley. aviad@eecs.berkeley.edu

†Princeton University. sethmw@cs.princeton.edu

# 1 Introduction

Consider a revenue-maximizing seller with  $n$  heterogeneous items for sale to a single buyer whose value for sets of items is unknown, but drawn from a known distribution  $D$ . When  $n = 1$ , seminal work of Myerson [18] and Riley and Zeckhauser [20] shows that the optimal selling scheme simply sets the price  $p^* = \arg \max\{p \cdot \Pr[v \geq p | v \leftarrow D]\}$ . Thirty years later, understanding the structure of the optimal mechanism when  $n > 1$  still remains a central open problem. Unfortunately, it is well-known that the optimal mechanism may require randomization, behave non-monotonically, and be computationally hard to find, even in very simple instances [22, 19, 4, 10, 9, 13, 14]. In light of this, recent work began studying the performance of especially simple auctions through the lens of approximation. Remarkably, these works have shown that when the bidder’s valuation is *additive*<sup>1</sup>, and her value for each item is drawn independently, very simple mechanisms can achieve quite good approximation ratios. Specifically, techniques developed in this series of works proves that the better of setting Myerson’s reserve on each item separately or setting Myerson’s reserve on the grand bundle of all items together achieves a 6-approximation [12, 17, 1].

While this model of buyer values is certainly mathematically interesting and economically motivated, it is also perhaps too simplistic to have broad real-world applications. A central question left open by these works is whether or not simple mechanisms can still approximate optimal ones in more general settings. In this work we resolve this question in the affirmative, showing that the better of selling separately (we will henceforth use SREV to denote the revenue of the optimal such mechanism) or together (henceforth BREV) still obtains a constant-factor approximation to the optimal revenue (henceforth REV) when buyer values are combinatorial in nature but complement-free.

*Informal Theorem 1.* Let  $D$  be any distribution over subadditive valuation functions with independent items. Then  $\max\{\text{SREV}, \text{BREV}\} \geq \Omega(1) \cdot \text{REV}$ . Furthermore, prices providing this guarantee can be found computationally efficiently.

We postpone a formal definition of exactly what it means for  $D$  to have “independent items” to Section 2. We note here a few instantiations of our model in commonly studied settings (from least to most general):

- $k$ -demand: The buyer has value  $v_i$  for item  $i$ , and the  $v_i$ s are drawn independently. The buyer’s value for a set  $S$  is  $v(S) = \max_{T \subseteq S, |T| \leq k} \{\sum_{i \in T} v_i\}$ .
- Additive up to constraints  $\mathcal{I}$ :  $\mathcal{I}$  is some downwards-closed set system on  $[n]$ . The buyer has value  $v_i$  for item  $i$ , and the  $v_i$ s are drawn independently.  $v(S) = \max_{T \subseteq S, T \in \mathcal{I}} \{\sum_{i \in T} v_i\}$ .
- Fractionally-subadditive: buyer has “possible values”  $\{v_{ij}\}_j$  for item  $i$ , and the sets  $\{v_{ij}\}_j$  are drawn independently across items (but may be correlated within an item).  $v(S) = \max_j \{\sum_{i \in S} v_{ij}\}$ .

## 1.1 Challenges of Combinatorial Valuations

The design of simple, approximately optimal mechanisms for any non-trivial multi-item setting has been a large focus for much of the Algorithmic Game Theory community over the past decade. Even “simple” settings with additive or unit-demand valuations required significant breakthroughs. The key insight enabling these breakthroughs for additive buyers is that the buyer’s valuation is separable across items. While the optimal mechanism can still be quite bizarre despite this realization [14], this fact enables certain elementary decomposition theorems that are surprisingly powerful (e.g. the “Marginal Mechanism” [12]). However, these

---

<sup>1</sup>A valuation function  $v(\cdot)$  is additive if  $v(S \cup T) = v(S) + v(T)$  for all  $S \cap T = \emptyset$ .

theorems are extremely sensitive to being able to separate the marginal contribution of different items *exactly* (and not just via upper/lower bounds). This is due to the phenomenon that a slight miscalculation in estimating a buyer’s value may cause her to change preferences entirely, resulting in a potentially unbounded loss of revenue. One of our main technical contributions is overcoming this obstacle by providing an *approximate* version of these decomposition theorems.

A further complication in applying these previous techniques is that they all make use of the fact that  $\text{SREV}(D_1 \times \dots \times D_n) = \sum_i \text{SREV}(D_i)$ . This claim is not even approximately true for subadditive buyers, and the ratio between the two values could be as large as  $n$  (the right-hand side is always larger). To have any hope of applying these tools, we therefore need a proxy for SREV that at least approximately has this separability property.

For unit-demand buyers, the key insight behind the mechanisms designed in [6, 7, 8, 16] is that every multi-dimensional problem instance has a related single-dimensional problem instance, and there is a correspondence between truthful mechanisms in the two instances. This realization means that one can instead design mechanisms for the single-dimensional setting, where optimal mechanisms are well understood due to Myerson’s virtual values, and translate them in a black-box manner to mechanisms for the original instance. While these techniques have proven extremely fruitful in the design of mechanisms for multiple unit-demand buyers and sophisticated feasibility constraints, they have also proven to be limited in use to unit-demand settings. A special case of our results can be seen as providing an alternative proof of the single-buyer result of Chawla, Hartline, and Kleinberg [6] (albeit with a significantly worse constant) that doesn’t require virtual valuation machinery.

Aside from the difficulties in applying existing machinery to design optimal mechanisms for combinatorial valuations, formal barriers exist as well. For instance, it is a trivial procedure for an additive buyer to select his utility-maximizing set of items when facing an item-pricing, and finding the revenue-optimal item-pricing is also trivial (just find the optimal price for each item separately). Yet for a subadditive buyer, both tasks are quite non-trivial. Just *computing* the expected revenue obtained by a fixed item-pricing is NP-hard. Worse, the buyer’s problem of just *selecting* her utility-maximizing set from a given item-pricing is also NP-hard! Therefore, buyers may behave quite unpredictably in the face of an item-pricing depending on how well they can optimize. Moreover, even if we are willing to assume that the buyer has the computational power to select her utility-maximizing set, it is known *still* that (without our independence assumption) finding an  $n^c$ -approximately optimal mechanism is NP-hard for all  $c = O(1)$  [5]. We sidestep all these difficulties by not attempting to compute or approximate SREV at all, nor trying to predict bizarre buyer behavior. We instead perform our analysis on revenue contributions only of items purchased *when the buyer is not willing to purchase any others*. Buyer behavior in such instances is predictable and easily computable: simply purchase the unique item for which  $v(\{i\}) > p_i$ . It is surprising that such an analysis suffices, as it completely ignores any revenue contribution coming from the entirely plausible event that the buyer is willing to purchase multiple items.

## 1.2 Techniques

We prove our main theorem by making use of the core-tail decomposition framework introduced by Li and Yao [17]. There are three high-level steps to applying the framework. The first is proving a “core decomposition” lemma that separates the optimal revenue into contributions from items which the buyer values very highly (the “tail”), and items which the buyer values not so high (the “core”). The second is showing that the contribution from the tail can be approximated well by SREV. The third is showing that the contribution from the core can be approximated well by  $\max\{\text{SREV}, \text{BREV}\}$ .

**The Core Decomposition Lemma.** The proof of the original Core Decomposition Lemma in [17] was obtained by cleverly stringing together simple claims proved in [12]. As discussed above,

these seemingly “obvious” claims may not extend beyond additive valuations over independent items, due to the fact that the buyer’s value cannot be separated across items. Nevertheless, we are able to prove an approximate version of the core decomposition lemma for subadditive buyers (Lemma 5) by making use of ideas from reductions from  $\epsilon$ -truthful mechanisms to fully truthful ones. Like in [1], our core decomposition lemma holds for many buyers. The proof for a single buyer, which is the focus of this paper can be found in Section 3.1. We also provide (Appendix E) a more technically involved proof for many buyers which builds on heavier tools from [3, 15, 11].

**Bounding the Tail’s Contribution.** Arguments for bounding the contribution from the tail in prior work (and ours) use the following reasoning. If the cutoff between core and tail is sufficiently high, then the probability that  $k$  items are simultaneously in the tail for a sampled valuation decays exponentially in  $k$ . If one can also show that the approximation guarantee of SREV decays subexponentially in  $k$ , then we can bound the gap between SREV and the tail’s contribution by a constant factor. We show that indeed the approximation guarantee of SREV decays only polynomially in  $k$ .

**Bounding the Core’s Contribution.** Arguments for bounding the contribution from the core in prior work (and ours) use the following reasoning. The total expected value for items in the core is a subadditive function of independent random variables (bounded above by the core-tail cutoff). If the cutoff between core and tail is sufficiently low, then one of two things must happen. Either the expected contribution from the core is also small, in which case SREV itself provides a good approximation, or the expected contribution is large, and therefore also large with respect to the cutoffs. In the latter case, a concentration bound implies that BREV must provide a good approximation. In the additive case, the appropriate concentration bound is Chebyshev’s inequality. In the subadditive case, we need heavier tools, and apply a concentration bound due to Schechtman [21].

### 1.3 Connection to Approximate Revenue-Monotonicity

Consider designing revenue-optimal mechanisms for two different markets, and suppose that the valuations of the consumers in the first market, first-order stochastically dominate<sup>2</sup> the valuations of the consumers in the second market. It then seems reasonable to expect that the optimal revenue achieved from the first market,  $\text{REV}(D^+)$ , should be at least as large as the revenue achieved from the dominated market,  $\text{REV}(D)$ . When there is just a single item for sale, this is an easy corollary of the format for Myerson’s optimal auction. Yet Hart and Reny provided an example where this intuition breaks even in a setting as simple as an *additive buyer with i.i.d. values for two items* [14]. Surprisingly, their example shows that it is possible to make *strictly* more revenue in a market when buyers have *strictly* less value for your goods, and the market need not even be very complex for this phenomenon to occur.

A natural question to ask then, is how large this anomaly can be. For example, Hart and Reny’s constructions exhibit a (multiplicative) gap of  $33/32$  between  $\text{REV}(D^+)$  and  $\text{REV}(D)$  for an additive buyer with correlated values for two items, and  $(1 + \frac{1}{7000000})$  for an additive buyer with i.i.d. values for two items. Interestingly, the simple mechanisms of [12, 17, 1] upper bound the possible gap of any instance where their results apply, since SREV and BREV are monotone for additive buyers (i.e.  $\text{SREV}(D^+) \geq \text{SREV}(D)$  and  $\text{BREV}(D^+) \geq \text{BREV}(D)$ ). Specifically, for an additive buyer the gap is at most  $(1 + 1/e)$  for two i.i.d. items, 2 for two asymmetric independent items, and 6 for any number of independent items. In Section 4 we show that as a corollary of our results, the gap is also constant for a subadditive buyer

---

<sup>2</sup>We say that a distribution  $D^+$  over valuation functions  $v^+$  first-order stochastically dominates distribution  $D$  over valuation functions  $v$  if the probability spaces can be coupled so that for every subset  $S$ ,  $v^+(S) \geq v(S)$ .

with independent items. Interestingly, this connection between approximately optimal simple mechanisms and approximate revenue-monotonicity is also fruitful in the other direction: it turns out that improving the bound on approximate monotonicity for a subadditive buyer would also improve the constant in our main theorem. Finally, we show in Subsection 4.3 that for an additive buyer with correlated values for two items, the gap is potentially infinite. (This is the case for which Hart and Reny provide a gap of  $33/32$ .) The proof is by a black-box reduction from an example due to Hart and Nisan [13] that exhibits a similar gap between simple and optimal mechanisms, further demonstrating the connection between these two important research directions.

## 1.4 Discussion and Open Problems

Our work contributes to the recent growing literature on simple, approximately optimal mechanisms. We extend greatly beyond prior work, providing the first simple and approximately optimal mechanisms for buyers with combinatorial valuations. Prior to our work, virtually nothing was known about this setting (modulo the impossibility result of [5]). Our results also demonstrate the strength of the core-tail decomposition framework developed by Li and Yao to go beyond additive buyers. We suspect that this framework will continue to prove useful in other Bayesian mechanism design problems.

In our opinion, the most exciting open question in this area is extending these results to multiple buyers. A beautiful lookahead reduction was recently developed by Yao [23] for additive buyers. Still, generalizing his tools beyond additive buyers seems quite challenging and is a very intriguing direction. Another important direction is extending our understanding of simple mechanisms to models of limited correlation over values for disjoint sets of items<sup>3</sup>. Recent independent work of Bateni et. al. [2] addresses this direction, providing approximation guarantees on  $\max\{\text{SREV}, \text{BREV}\}$  vs.  $\text{REV}$  for an additive buyer whose values for items are drawn from a common-base-value distribution and various extensions, and their results also make use of a core-tail decomposition. A natural question in this direction is whether our results extend to settings where buyer values are both combinatorial *and* exhibit limited correlation between disjoint sets of items, as the end goal is to have a model that encompasses as many real-world instances as possible.

## 2 Preliminaries

We focus the body of the exposition on the single-buyer problem, and defer all details regarding auctions for multiple buyers to Appendix E. There is a single revenue-maximizing seller with  $n$  items for sale to a single buyer. The buyer has combinatorial valuations for the items (i.e. value  $v(S)$  for receiving set  $S$ ), and is quasi-linear and risk-neutral. That is, the buyer’s utility for a randomized outcome that awards him set  $S$  with probability  $A(S)$  while paying (expected) price  $p$  is  $\sum_S A(S) v(S) - p$ .  $v(\cdot)$  is unknown to the seller, who has a prior  $D$  over possible buyer valuations that is *subadditive over independent items*, a term we describe below. By the taxation principle, the seller may restrict attention to only lottery systems. In other words, the seller provides a list of potential lotteries (distributions over sets) each with a price, and the buyer chooses the utility-maximizing option.

### 2.1 Subadditive valuations over independent items

We now carefully define what we mean by subadditive valuations over independent items. Intuitively, our model is such that the buyer has some private information  $x_i$  pertaining to item  $i$ ,<sup>4</sup> and  $D^{\vec{x}}$  is a product distribution over  $\mathbb{R}^n$  representing the seller’s prior over the private

<sup>3</sup>Note that as we mention in the previous section, for arbitrary correlated items the gap can be infinite [13, 4].

<sup>4</sup>Think of this information as “information about the buyer’s preferences related to item.”

information possessed by the buyer. The buyer's valuation for set  $S$  is parametrized by the private information she has about items in that set, and can be written as  $V(\{x_i\}_{i \in S}, S)$ . In economic terms, this models that the items not received by the buyer impose no externalities. We capture this formally in the definition below:

**Definition 1.** We say that a distribution  $D$  over valuation functions  $v(\cdot)$  is subadditive over independent items if:

1. All  $v(\cdot)$  in the support of  $D$  exhibit *no externalities*. That is, there exists a distribution  $D^{\vec{x}}$  over  $\mathbb{R}^n$  and a function  $V$  such that  $D$  is the distribution that first samples  $\vec{x} \leftarrow D^{\vec{x}}$  and outputs the valuation function  $v(\cdot)$  with  $v(S) = V(\{x_i\}_{i \in S}, S)$  for all  $S$ .
2. All  $v(\cdot)$  in the support of  $D$  are *monotone*. That is,  $v(S) \leq v(S \cup T)$  for all  $S, T$ .
3. All  $v(\cdot)$  in the support of  $D$  are *subadditive*. That is,  $v(S \cup T) \leq v(S) + v(T)$ .
4. The private information is *independent across items*. That is, the  $D^{\vec{x}}$  guaranteed in property 1. is a product distribution.

We describe now how to encode commonly studied valuation distributions in this model.

**Example.** The following types of distributions can be modeled as subadditive over independent items.

1. Additive:  $x_i$  is the buyer's value for item  $i$ .  $V(\{x_i\}_{i \in S}, S) = \sum_{i \in S} x_i$ .
2.  $k$ -demand:  $x_i$  is the buyer's value for item  $i$ .  $V(\{x_i\}_{i \in S}, S) = \max_{T \subseteq S, |T| \leq k} \{\sum_{i \in T} x_i\}$ .
3. Additive up to  $\mathcal{I}$ :  $x_i$  is the buyer's value for item  $i$ .  $V(\{x_i\}_{i \in S}, S) = \max_{T \subseteq S, T \in \mathcal{I}} \{\sum_{i \in T} x_i\}$ .
4. Fractionally subadditive:  $x_i$  encodes (using bit interleaving, for instance) the values  $\{v_{ij}\}_j$ .  $V(\{x_i\}_{i \in S}, S) = \max_j \{\sum_{i \in S} v_{ij}\}$ .

## 2.2 Notation

**Definition 2.** For any distribution  $D$  of buyer's valuation, we use the following notation, most of which is due to [12, 1]:

- $D_i$ : The distribution of  $v(\{i\})$  when  $v(\cdot) \leftarrow D$ .
- $t$ : the cutoff between core and tail. If  $v(\{i\}) > t$ , we say that item  $i$  is in the tail. Otherwise it is in the core.
- $D_A$ : the distribution  $D$ , conditioned on  $A$  being exactly the set of items in the tail.
- $D_A^T$ : the distribution  $D_A$  restricted just to items in the tail (i.e.  $A$ ).
- $D_A^C$ : the distribution  $D_A$  restricted just to items in the core (i.e.  $\bar{A}$ ).
- $p_i$ : the probability that element  $i$  is in the tail.
- $p_A$ : the probability that  $A$  is exactly the set of items in the tail (note that  $p_{\{i\}} \neq p_i$ ).
- $\text{REV}(D)$ : The maximum revenue obtainable via a truthful mechanism from a buyer with valuation profile drawn from  $D$ .
- $\text{BREV}(D)$ : The revenue obtainable from a buyer with valuation profile drawn from  $D$  by auctioning the grand bundle via Myerson's optimal auction.

- $\text{SREV}(D)$ : The maximum revenue obtainable from a buyer with valuation profile drawn from  $D$  by pricing each item separately. Note that when the buyer is not additive,  $\text{SREV}(D)$  behaves erratically and is NP-hard to find [9].
- $\text{REV}_q(D)$ : For a one-dimensional distribution  $D$ , the optimal revenue obtained by a reserve price that sells with probability at most  $q$ .<sup>5</sup>
- $\text{SREV}_{\vec{q}}^*(D) : \prod_i (1 - q_i) \cdot \sum_i \text{REV}_{q_i}(D_i)$ : a proxy for  $\text{SREV}(D)$  that behaves nicer and is easy to compute.
- $\text{VAL}(D)$ : the buyer’s expected valuation for the grand bundle,  $\mathbb{E}_{v \leftarrow D} [v([n])]$ .

When the distribution is clear from the context, we simply use  $\text{REV}$ ,  $\text{VAL}$ , etc. Most of this notation is standard following [12], with the exception of  $\text{REV}_q$  and  $\text{SREV}_{\vec{q}}^*$ . We introduce  $\text{SREV}_{\vec{q}}^*$  because it will serve as a proxy to  $\text{SREV}$  that behaves nicely and is easy to compute. Note that  $\text{SREV}_{\vec{q}}^*$  is essentially computing the revenue of the best item pricing that sells item  $i$  with probability at most  $q_i$ , but only counting revenue from cases where the other values are too low to have possibly sold (and actually it undercounts this quantity). We conclude the preliminaries by stating a lemma of Hart and Nisan that we will use. We include the proof in the appendix for completeness, as well as to verify that it continues to hold when the valuations are not additive.

**Lemma 1** (special case of Sub-domain Stitching [12]).  $\text{REV}(D) \leq \sum_A p_A \text{REV}(D_A)$ .

### 3 Main Result: Constant-factor approximation for subadditive buyer

**Theorem 1.** *When  $D$  is subadditive over independent items, there exists a probability vector  $\vec{q}$  such that:*

$$\text{REV}(D) \leq 314 \text{SREV}_{\vec{q}}^*(D) + 24 \text{BREV}(D)$$

*Furthermore,  $\vec{q}$  can be computed efficiently, as well as an induced item pricing that yields expected revenue at least  $\text{SREV}_{\vec{q}}^*(D)$ .*

#### Proof outline

We follow the core-tail decomposition framework. First, we provide an approximate core decomposition lemma in Section 3.1. Then, we provide a bound on the contribution of the core with respect to  $\max\{\text{BREV}, \text{SREV}^*\}$  in Section 3.2, and a bound on the contribution of the tail with respect to  $\text{SREV}^*$  as a function of the cutoffs chosen in Section 3.3.

For ease of exposition, we simply set  $t$  so that the probability of having an empty tail is exactly half; i.e.  $p_\emptyset = \prod (1 - p_i) = 1/2$ . We also set  $\vec{q} = \vec{p}$ .

#### 3.1 Approximate Core Decomposition

In this section, we prove our approximate core decomposition lemma. The key ingredient will be an approximate version of the “Marginal Mechanism” lemma from [12] for subadditive buyers, stated below:

---

<sup>5</sup>We assume, w.l.o.g. that for all one-dimensional  $D$  and all  $q \in [0, 1]$  it is possible to set a price that sells with probability exactly  $q$ . Formally, we achieve this by adding an infinitesimally small tie-breaker sampled from  $[0, dx]$  to every real number. For instance, if we wanted a reserve that sold with probability  $1/2$  to the distribution that is a point mass at 1, we would set price  $1 + dx/2$ .

**Lemma 2.** (“Approximate Marginal Mechanism”)

Let  $S, T$  be a partition of  $[n]$ , and let  $D = (D_S, D_T)$  be the joint distribution for the valuations of items in  $S, T$ , respectively, for buyers with subadditive valuations. Then for any  $0 < \epsilon < 1$ ,

$$\text{REV}(D) \leq \left( \frac{1}{\epsilon} + \frac{1}{1-\epsilon} \right) \text{VAL}(D_S) + \frac{1}{1-\epsilon} \mathbb{E}_{v_S \leftarrow D_S} [\text{REV}(D_T | v_S)] \quad (1)$$

When  $D_S$  and  $D_T$  are independent, this simplifies to

$$\text{REV}(D) \leq \left( \frac{1}{\epsilon} + \frac{1}{1-\epsilon} \right) \text{VAL}(D_S) + \frac{1}{1-\epsilon} \text{REV}(D_T).$$

We outline the proof of Lemma 2 below. We first recall Hart and Nisan’s [12] original “Marginal Mechanism” lemma (that holds for an additive buyer without any multipliers). We provide a complete proof so that the reader can see where the argument fails for subadditive buyers.

**Lemma 3** (“Marginal Mechanism” [12]). Let  $S \sqcup T$  be any partition of  $[n]$ , and let  $D^+$  be any distribution over valuation functions such that  $v^+(U) = v^+(U \cap S) + v^+(U \cap T)$  for all  $U \subseteq n$ , and  $v^+$  in the support of  $D^+$ . Let also  $D_S$  denote  $D^+$  restricted to items in  $S$  and  $D_T$  denote  $D^+$  restricted to items in  $T$ . Then  $\text{REV}(D^+) \leq \text{VAL}(D_S) + \mathbb{E}_{v_S \leftarrow D_S} [\text{REV}(D_T | D_S)]$ .

*Proof.* We design a mechanism  $M_T$  (the “Marginal Mechanism”) to sell items in  $T$  to consumers sampled from  $D_T | v_S$  based on the optimal mechanism  $M$  for selling items in  $S \sqcup T$  to consumers sampled from  $D^+$ . Define  $A(v)$  to be the (possibly random) allocation of items awarded to type  $v$  in  $M$ , and  $p(v)$  to be the price paid. Let  $M_T$  first sample a value  $v_S \leftarrow D_S$ . The buyer is then invited to report any type  $v_T$ , and  $M_T$  will award him the items in  $A(v_S, v_T) \cap T$  and charge him price  $p(v_S, v_T) - v_S(A(v_S, v_T) \cap S)$ . In other words, the buyer will receive value from exactly the same items in  $M_T$  as he would have received in  $M$ , except he receives the actual items in  $T$ , whereas for items in  $S$  he is given a monetary rebate instead of his actual value.

We first claim that if  $M$  is truthful, then so is  $M_T$ . The utility of a buyer with type  $v_T$  for reporting  $v'_T$  to  $M_T$  can be written as:  $v_T(A(v_S, v'_T) \cap T) + v_S(A(v_S, v'_T) \cap S) - p(v_S, v'_T) = (v_S, v_T)(A(v_S, v'_T)) - p(v_S, v'_T)$ , which is exactly the utility of a buyer with type  $(v_S, v_T)$  for reporting  $(v_S, v'_T)$  to  $M$ . As  $M$  was truthful, we know that a buyer with type  $(v_S, v_T)$  maximizes utility when reporting  $(v_S, v_T)$  over all possible  $(v_S, v'_T)$ . Therefore, a buyer with type  $v_T$  prefers to tell the truth as well.

Finally, we just have to compute the revenue of  $M_T$ . For each  $v_S$ , the marginal mechanism provides a concrete mechanism for the distribution  $D_T | v_S$  that attains revenue at least  $\text{REV}(D^+ | v_S) - v_S(S)$ . So  $\text{REV}(D_T | v_S) \geq \text{REV}(D^+ | v_S) - v_S(S)$ . Taking an expectation over all  $v_S$  (and an application of sub-domain stitching) yields the lemma.  $\square$

Notice that it is *crucial* in the proof above that the buyer’s value could be written as  $v_S(\cdot) + v_T(\cdot)$ . Otherwise the auctioneer does not know how much to “reimburse” the buyer, since the correct amount depends on the buyer’s private information. The buyer can then manipulate his own report  $v_T$  to influence how much he gets reimbursed for the items in  $S$ .

A natural approach then, given any distribution  $D$  over subadditive valuations, is to define a new value distribution  $D^+$  by redefining all  $v(\cdot)$  to satisfy  $v(U) = v(U \cap S) + v(U \cap T)$  (it is easy to see that all valuations in the support of  $D^+$  are still subadditive). Unfortunately, even though  $D^+$  (first-order stochastically) dominates  $D$ , due to non-monotonicity we could very well have  $\text{REV}(D^+) < \text{REV}(D)$ . Still, we bound the revenue lost as we move from  $D$  to  $D^+$  by making use of tools for turning  $\epsilon$ -truthful mechanisms into truly truthful ones. Lemma 4 and Corollary 1 below capture this formally. Both proofs appear in Appendix B.

**Lemma 4.** Consider two coupled distributions  $D$  and  $D^+$ , with  $v(\cdot)$  and  $v^+(\cdot)$  denoting a random sample from each. Define the random function  $\delta(\cdot)$  so that  $\delta(S) = v^+(S) - v(S)$  for all  $S$ . Suppose that  $\delta(S) \geq 0$  for all  $S$  and that  $\delta(\cdot)$  also satisfies  $\mathbf{E}_D [\max_{S \subseteq [n]} \{\delta(S)\}] \leq \bar{\delta}$ . Then for any  $0 < \epsilon < 1$ ,

$$\text{REV}(D^+) \geq (1 - \epsilon) (\text{REV}(D) - \bar{\delta}/\epsilon).$$

**Corollary 1.** For a given partition of  $[n]$ ,  $S \sqcup T$ , and distribution  $D$  over subadditive valuations, define  $D_S$  to be  $D$  restricted to items in  $S$ , and  $D^+$  to first sample  $v \leftarrow D$ , and output  $v^+(\cdot)$  with  $v^+(U) = v(U \cap S) + v(U \cap T)$ . Then for all  $\epsilon \in (0, 1)$ ,  $\text{REV}(D) \leq \frac{\text{REV}(D^+)}{1 - \epsilon} + \frac{\text{VAL}(D_S)}{\epsilon}$ .

The proof of Lemma 2 is now a combination of Corollary 1 and Lemma 3. We can now provide our approximate core decomposition by combing sub domain stitching (1) and approximate marginal mechanism (Lemma 2). The proof of Lemma 5 appears in Appendix B.

**Lemma 5.** (“Approximate Core Decomposition”)

For any distribution  $D$  that is subadditive over independent items, and any  $0 < \epsilon < 1$ ,

$$\text{REV}(D) \leq \left( \frac{1}{\epsilon} + \frac{1}{1 - \epsilon} \right) \text{VAL}(D_\emptyset^C) + \frac{1}{1 - \epsilon} \sum_{A \subseteq [n]} p_A \text{REV}(D_A^T).$$

In particular, for  $\epsilon = 1/2$ , we have

$$\text{REV}(D) \leq 4 \text{VAL}(D_\emptyset^C) + 2 \sum_{A \subseteq [n]} p_A \text{REV}(D_A^T)$$

### 3.2 Core

Here, we show how to bound  $\text{VAL}(D_\emptyset^C)$  using  $\max\{\text{SREV}_q^*(D), \text{BREV}(D)\}$ . We use a concentration result due to Schechtman [21] that first requires a definition.

**Definition 3.** Let  $D^{\vec{x}}$  denote a distribution of private information,  $V$  denote a valuation function  $V(\vec{x}, \cdot)$ , and  $D$  denote the distribution that samples  $\vec{x} \leftarrow D^{\vec{x}}$  and outputs the function  $v(\cdot) = V(\vec{x}, \cdot)$ . Then  $D$  is  $c$ -Lipschitz if for all  $\vec{x}, \vec{y}$ , and sets  $S$  and  $T$  we have:

$$|V(\vec{x}, S) - V(\vec{y}, T)| \leq c \cdot (|S \cup T| - |S \cap T| + |\{i \in S \cap T : x_i \neq y_i\}|)$$

Before applying Schechtman’s theorem, we show that  $D_\emptyset^C$  is  $t$ -Lipschitz (recall that  $t$  is the cutoff between core and tail). The proof of Lemma 6 appears in Appendix C.

**Lemma 6.** Let  $D$  be any distribution that is subadditive over independent items where each  $v(\{i\}) \in [0, c]$  with probability 1. Then  $D$  is  $c$ -Lipschitz.

**Corollary 2.**  $D_\emptyset^C$  is  $t$ -Lipschitz.

Now we state Schechtman’s theorem and apply it to bound  $\text{VAL}(D_\emptyset^C)$ .

**Theorem 2.** ([21]) Suppose that  $D$  is a distribution that is subadditive over independent items and  $c$ -Lipschitz. Then for any parameters  $q, a, k > 0$ ,

$$\Pr_{v \leftarrow D} [v([n]) \geq (q + 1)a + k \cdot c] \leq \Pr [v([n]) \leq a]^{-q} q^{-k}$$

In particular, if  $a$  is the median of  $v([n]) \mid_{v \leftarrow D}$  and  $q = 2$ , we have

$$\Pr_{v \leftarrow D} [v([n]) \geq 3a + k \cdot c] \leq 4 \cdot 2^{-k}$$

**Corollary 3.** *Suppose that  $D$  is a distribution that is subadditive over independent items and  $c$ -Lipschitz. If  $a$  is the median of  $v([n])$   $|_{v \leftarrow D}$ , then  $\mathbb{E}_{v \leftarrow D} [v([n])] \leq 3a + 4c/\ln 2$ .*

*Proof.*  $\mathbb{E}[v([n])] = \int_0^\infty \Pr[v([n]) > y] dy$ . We can upper bound this using the minimum of 1 and the bound provided in Theorem 2 to yield:

$$\mathbb{E}[v([n])] \leq 3a + \int_0^\infty 4 \cdot 2^{-y/c} dy = 3a + 4c/\ln 2$$

□

**Proposition 1.**  $\text{VAL}(D_\emptyset^C) \leq 6\text{BREV} + 4t/\ln 2$ .

*Proof.* Since  $a$  is the median of  $v([n])$ , we can set price  $a$  on the grand bundle and extract revenue at least  $a/2$ . Therefore,  $\text{BREV} \geq a/2$ . The proposition then follows by combining Corollaries 2 and 3. □

Finally, if the cutoff  $t$  is not too large, we can recover a constant fraction of it by selling each item separately. In particular, in Appendix C, we prove (recall that we set  $t$  to induce  $p_\emptyset = 1/2$ ):

**Lemma 7.**  $\text{SREV}_p^* \geq t/4$ .

Therefore,

$$\text{VAL}(D_\emptyset^C) \leq 6\text{BREV} + 24\text{SREV}_p^*.$$

### 3.3 Tail

We now show that the revenue from the tail can be approximated by  $\text{SREV}_{\tilde{q}}^*$ . We begin by proving a much weaker bound on the optimum revenue for any distribution of subadditive valuations over independent items:

**Lemma 8.**

$$\text{REV}(D) \leq 6n^{\log_2 6} \sum_i \text{REV}(D_i).$$

The proof recursively halves the number of items. The recursion is of depth  $\lceil \log_2 n \rceil$ , and each step loses a factor of 6 due to the approximate marginal mechanism (Lemma 2). The details appear in Appendix D.

Note that in Lemma 8,  $\sum_i \text{REV}(D_i)$  is *not* the same as  $\text{SREV}(D)$  as the buyer is not necessarily additive. In fact, they can be off by a factor of  $n$  (in the case of a unit-demand buyer). Nonetheless, this weak bound suffices for our analysis of the tail, which is summarized in Proposition 2 below. Essentially, the proposition amplifies the bound in Lemma 8 greatly by making use of the fact that it is unlikely to see multiple items in the tail.

**Proposition 2.**

$$\sum_A p_A \text{REV}(D_A^T) \leq 109 \cdot \text{SREV}_p^*$$

The proof of Proposition 2 appears in Appendix D. Note that Theorem 1 is now a corollary of Proposition 2, Lemma 7, and Lemma 5 (setting  $\epsilon = 1/2$ ).

## 4 Simple Auctions and Approximate Revenue Monotonicity

In this section we explore the rich connection between approximately optimal simple auctions, and approximate revenue monotonicity. By approximate revenue monotonicity, we formally mean the following:

**Definition 4.** We say that a class of distributions is  $\alpha$ -approximately revenue monotone if for any two distributions  $D$  and  $D^+$  in that class such that  $D^+$  first-order stochastically dominates  $D$ ,<sup>6</sup>  $\alpha \cdot \text{REV}(D^+) \geq \text{REV}(D)$ .

In the rest of the section we observe that subadditive valuations over independent items are  $\alpha$ -approximately monotone for some constant factor (Subsection 4.1). We also note that a (significantly) tighter approximate monotonicity would yield a better factor of approximation in Theorem 1 (Subsection 4.2). Finally, for the class of (possibly correlated) additive valuations over  $n$  items, we prove a reduction from approximate revenue monotonicity to approximately optimal simple auctions (that loses a factor of  $n$ ). Then we use an infinite gap between  $\max\{\text{BREV}, \text{SREV}\}$  and  $\text{REV}$  for two items due to Hart and Nisan [13] to prove an infinite lower bound on approximate revenue monotonicity (Subsection 4.3).

### 4.1 Approximately optimal simple auctions imply approximate monotonicity

As a corollary of our main theorem (Theorem 1) we deduce constant-factor approximate monotonicity for subadditive valuations over independent items:

**Corollary 4.** *The class of subadditive valuations over independent items is 338-approximately monotone.*

Similarly, from the 6-approximation of Babaioff et al. for additive yields.

**Corollary 5.** *The class of additive valuations over independent items is 6-approximately monotone.*

*Proof.* For additive functions,  $\text{BREV}$  and  $\text{SREV}$  constitute of separate Myerson's auctions, and are therefore revenue monotone. Thus,

$$\begin{aligned} 6\text{REV}(D^+) &\geq 6 \max\{\text{BREV}(D^+), \text{SREV}(D^+)\} \\ &\geq 6 \max\{\text{BREV}(D), \text{SREV}(D)\} \geq \text{REV}(D) \end{aligned}$$

For subadditive functions,  $\text{SREV}(D)$  is no longer monotone, but  $\text{SREV}_{\bar{q}}^*(D)$  is. This is because  $\text{SREV}_{q_i}(D_i)$  is clearly monotone, and  $\text{SREV}_{\bar{q}}^*$  is just a (scaled) sum of  $\text{SREV}_{q_i}(D_i)$ . So we get that there exists a  $\bar{q}$  such that:

$$\begin{aligned} 338\text{REV}(D^+) &\geq 338 \max\{\text{BREV}(D^+), \text{SREV}_{\bar{q}}^*(D^+)\} \\ &\geq 338 \max\{\text{BREV}(D), \text{SREV}_{\bar{q}}^*(D)\} \geq \text{REV}(D) \end{aligned}$$

□

### 4.2 Approximate monotonicity implies approximately optimal simple auctions

A closer look at the proof of our main theorem also yields the converse of the above corollaries, namely: a tighter approximate monotonicity for subadditive valuations would yield an improved factor of approximation by simple auctions, as well as a simpler proof.

<sup>6</sup>Recall that we say  $D^+$  first-order stochastically dominates  $D$  if they can be coupled so that when we sample  $v^+$  from  $D^+$  and  $v$  from  $D$  we have  $v^+(S) \geq v(S)$  for all  $S$ .

**Corollary 6.** *If the class of subadditive valuations over independent items is  $\alpha$ -approximately monotone, then*

$$\text{REV} \leq \alpha \left( (37\alpha + 24) \text{SREV} + 6\text{BREV} \right)$$

*Proof.* (Sketch)

Recall that in the proof of the Approximate Marginal Mechanism Lemma (Lemma 2), we made use of Lemma 4 to bound the gap between  $\text{REV}(D^+)$  and  $\text{REV}(D)$ , where  $D^+$  first-order stochastically dominated  $D$ . Instead of the  $\epsilon$ -truthful to truthful reduction, we could derive  $\alpha \text{REV}(D^+) \geq \text{REV}(D)$  from approximate monotonicity. Then, we can directly apply Lemma 3 to get:

$$\text{REV}(D) \leq \alpha (\text{VAL}(D_S) + \mathbb{E}_{v_S \sim D_S} [\text{REV}(D_T \mid \mathbf{v}_S)])$$

Instead of

$$\text{REV}(D) \leq 4\text{VAL}(D_S) + 2\mathbb{E}_{v_S \sim D_S} [\text{REV}(D_T \mid \mathbf{v}_S)]$$

If  $\alpha \leq 2$ , this indeed yields a tighter approximation.  $\square$

### 4.3 Correlated Distributions are not Approximately Monotone

So far we've shown that (for some valuation classes) approximately optimal simple mechanisms imply approximate revenue-monotonicity. Are all classes approximately revenue-monotone? In this subsection we provide a reduction from an instance where  $\max\{\text{SREV}, \text{BREV}\}$  does not approximate  $\text{REV}$  to show an infinite non-monotonicity for correlated items. We first prove a reduction from gaps between  $\text{BREV}$  and  $\text{REV}$  to non-monotonicity.

**Proposition 3.** *Let  $D$  be a distribution over subadditive valuations for  $n$  items for which  $\text{REV}(D) > c \cdot \text{BREV}(D)$ . Then any class of distributions containing  $D$  and all single-dimensional distributions<sup>7</sup> is not  $(c/n)$ -approximately revenue monotone.*

*Proof.* We define  $D^+$  as follows. First, sample  $v \leftarrow D$ . Then let  $i^* = \arg \max_i \{v(\{i\})\}$ . Now, set  $v^+(S) = v(\{i^*\}) \cdot |S|$  for all  $|S|$ . By subadditivity, it's clear that  $D^+$  first order stochastically dominates  $D$ . Now, however,  $D^+$  is a single-dimensional distribution, meaning that  $\text{BREV}(D^+) = \text{REV}(D^+)$  [18, 20]. Finally, we just need to compare  $\text{BREV}(D)$  to  $\text{BREV}(D^+)$ .

Note that by monotonicity, we have  $v^+([n]) \leq n \cdot v([n])$  for all  $v, v^+$ . Therefore, for any price  $p$ , if  $v^+([n]) > p$ ,  $v([n]) > p/n$ . This immediately implies that  $\text{BREV}(D) \geq \text{BREV}(D^+)/n$ : let  $p$  be the optimal reserve for the grand bundle under  $D^+$ , then setting price  $p/n$  sells with at least the same probability under  $D$ . Putting both observations together, we see that:  $\text{REV}(D) > c \text{BREV}(D) \geq (c/n) \text{BREV}(D^+) = (c/n) \text{REV}(D^+)$ , meaning that any class containing  $D$  and  $D^+$  is not  $(c/n)$ -approximately monotone.  $\square$

We apply Proposition 3 to a theorem of Hart and Nisan.

**Theorem 3.** (Hart and Nisan [13]) *There exists a distribution  $D$  over correlated additive valuations for two items such that  $\text{BREV}(D) \leq 1/2$ , and  $\text{REV}(D) = \infty$ .*

**Corollary 7.** *There exist distributions  $D$  and  $D^+$  over correlated additive valuations for two items such that  $D^+$  first-order stochastically dominates  $D$ ,  $\text{REV}(D^+) = 1$ , and yet  $\text{REV}(D) = \infty$ . Therefore, the class of correlated additive valuations for two items is not  $c$ -approximately revenue monotone for any finite  $c$ .*

*Acknowledgement.* We would like to thank Moshe Babaioff, Hu Fu, Nicole Immorlica, and Brendan Lucier for numerous suggestions and helpful discussions.

<sup>7</sup>A distribution is single-dimensional if for all  $v$  in its support,  $v(S) = c|S|$  for some value  $c$ .

## References

- [1] Moshe Babaioff, Nicole Immorlica, Brendan Lucier, and S. Matthew Weinberg. A simple and approximately optimal mechanism for an additive buyer. *FOCS*, 2014. (document), 1, 1.2, 1.3, 2, D, E.1
- [2] MohammadHossein Bateni, Sina Dehghani, MohammadTaghi Hajiaghayi, and Saeed Seddighin. Revenue Maximization for Selling Multiple Correlated Items. *arxiv report*, 2015. <http://arxiv.org/abs/1412.3187>. 1.4
- [3] Xiaohui Bei and Zhiyi Huang. Bayesian Incentive Compatibility via Fractional Assignments. In *the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2011. 1.2, E, E.3
- [4] Patrick Briest, Shuchi Chawla, Robert Kleinberg, and S. Matthew Weinberg. Pricing randomized allocations. In *SODA*, pages 585–597, 2010. 1, 3
- [5] Yang Cai, Constantinos Daskalakis, and S. Matthew Weinberg. Understanding Incentives: Mechanism Design becomes Algorithm Design. In *the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2013. <http://arxiv.org/pdf/1305.4002v1.pdf>. 1.1, 1.4
- [6] Shuchi Chawla, Jason D. Hartline, and Robert Kleinberg. Algorithmic pricing via virtual valuations. In *Proceedings of the 8th ACM Conference on Electronic Commerce, EC '07*, pages 243–251, 2007. 1.1
- [7] Shuchi Chawla, Jason D. Hartline, David L. Malec, and Balasubramanian Sivan. Multi-Parameter Mechanism Design and Sequential Posted Pricing. In *the 42nd ACM Symposium on Theory of Computing (STOC)*, 2010. 1.1
- [8] Shuchi Chawla, David L. Malec, and Balasubramanian Sivan. The Power of Randomness in Bayesian Optimal Mechanism Design. In *the 11th ACM Conference on Electronic Commerce (EC)*, 2010. 1.1
- [9] Xi Chen, Ilias Diakonikolas, Dimitris Paparas, Xiaorui Sun, and Mihalis Yannakakis. The complexity of optimal multidimensional pricing. In *the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2014. 1, 2
- [10] Constantinos Daskalakis, Alan Deckelbaum, and Christos Tzamos. The complexity of optimal mechanism design. In *the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2014. 1
- [11] Constantinos Daskalakis and S. Matthew Weinberg. Symmetries and Optimal Multi-Dimensional Mechanism Design. In *the 13th ACM Conference on Electronic Commerce (EC)*, 2012. 1.2, E, E.2, E.3, 4, 13
- [12] Sergiu Hart and Noam Nisan. Approximate revenue maximization with multiple items. In *Proceedings of the 13th ACM Conference on Electronic Commerce, EC '12*, pages 656–656, 2012. 1, 1.1, 1.2, 1.3, 2, 1, 3.1, 3.1, 3, 1, D
- [13] Sergiu Hart and Noam Nisan. The menu-size complexity of auctions. In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce, EC '13*, pages 565–566, 2013. 1, 1.3, 3, 4, 3
- [14] Sergiu Hart and Philip J. Reny. Maximal Revenue with Multiple Goods: Nonmonotonicity and Other Observations. Discussion Paper Series dp630, The Center for the Study of Rationality, Hebrew University, Jerusalem, November 2012. (document), 1, 1.1, 1.3

- [15] Jason D. Hartline, Robert Kleinberg, and Azarakhsh Malekian. Bayesian Incentive Compatibility via Matchings. In *the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2011. [1.2](#), [E](#), [E.3](#), [10](#), [8](#), [E.3](#), [11](#), [12](#)
- [16] Robert Kleinberg and S. Matthew Weinberg. Matroid Prophet Inequalities. In *the 44th Annual ACM Symposium on Theory of Computing (STOC)*, 2012. [1.1](#)
- [17] Xinye Li and Andrew Chi-Chih Yao. On revenue maximization for selling multiple independently distributed items. *Proceedings of the National Academy of Sciences*, 110(28):11232–11237, 2013. [\(document\)](#), [1](#), [1.2](#), [1.2](#), [1.3](#)
- [18] Roger B. Myerson. Optimal Auction Design. *Mathematics of Operations Research*, 6(1):58–73, 1981. [1](#), [4.3](#)
- [19] Gregory Pavlov. A Property of Solutions to Linear Monopoly Problems. *B. E. Journal of Theoretical Economics*, 11(1), 2011. Article 4. [1](#)
- [20] John Riley and Richard Zeckhauser. Optimal Selling Strategies: When to Haggle, When to Hold Firm. *Quarterly J. Economics*, 98(2):267–289, 1983. [1](#), [4.3](#)
- [21] Gideon Schechtman. Concentration, results and applications, 1999. [1.2](#), [3.2](#), [2](#)
- [22] John Thanassoulis. Hagglng Over Substitutes. *J. Economic Theory*, 117:217–245, 2004. [1](#)
- [23] Andrew Chi-Chih Yao. An n-to-1 bidder reduction for multi-item auctions and its applications. In *To Appear in the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2015. [1.4](#)

## A Sub-domain stitching

The following lemma is due to Hart and Nisan. We reproduce its proof below for completeness and to verify that it continues to hold for subadditive (in fact, arbitrary) valuations.

**Lemma 1** (special case of Sub-domain Stitching [\[12\]](#)).  $\text{REV}(D) \leq \sum_A p_A \text{REV}(D_A)$ .

*Proof.* Let  $M$  be an optimal mechanism for selling items with valuations sampled from  $D$ , and let  $\text{REV}_M(D) = \text{REV}(D)$  denote its revenue. Then,  $\text{REV}_M(D) = \sum p_A \text{REV}_M(D_A)$ . Also, for each  $A \subseteq [n]$ ,  $\text{REV}_M(D_A) \leq \text{REV}(D_A)$ .  $\square$

## B Omitted Proofs from Section [3.1](#).

**Lemma 4.** *Consider two coupled distributions  $D$  and  $D^+$ , with  $v(\cdot)$  and  $v^+(\cdot)$  denoting a random sample from each. Define the random function  $\delta(\cdot)$  so that  $\delta(S) = v^+(S) - v(S)$  for all  $S$ . Suppose that  $\delta(S) \geq 0$  for all  $S$  and that  $\delta(\cdot)$  also satisfies  $\mathbf{E}_D [\max_{S \subseteq [n]} \{\delta(S)\}] \leq \bar{\delta}$ . Then for any  $0 < \epsilon < 1$ ,*

$$\text{REV}(D^+) \geq (1 - \epsilon) (\text{REV}(D) - \bar{\delta}/\epsilon).$$

*Proof.* Consider a mechanism  $M$  which achieves the optimal revenue  $\text{REV}(D)$ . Let  $(\phi_v, p_v)$  denote the lottery purchased by a buyer with type  $v$  in  $M$ , where  $\phi_v$  is a (possibly randomized) allocation, and  $p$  is a price. Consider now the mechanism  $M'$  that offers the same menu as  $M$ , but with all prices discounted by a factor of  $(1 - \epsilon)$ . Let  $(\phi'_v, p'_v)$  denote the lottery that a buyer

with type  $v'$  (coupled with  $v$ ) chooses to purchase in  $M'$ . The following inequalities must hold (we will abuse notation and let  $v(A) = \mathbf{E}_{S \leftarrow A}[v(S)]$ ):

$$v(\phi_v) - p_v \geq v(\phi'_v) - p'_v. \quad (2)$$

$$v'(\phi'_v) - (1 - \epsilon)p'_v \geq v'(\phi_v) - (1 - \epsilon)p \quad (3)$$

Now, summing equations (2) and (3) (then making use of the definition of  $\delta(\cdot)$  and the fact that it is non-negative), we have:

$$\begin{aligned} \epsilon p'_v + \delta(\phi'_v) &\geq \epsilon p \\ \Rightarrow p' &\geq p - \delta(\phi'_v)/\epsilon \end{aligned}$$

We can now bound the expected revenue by taking an expectation over all valuations:

$$\begin{aligned} \text{REV}(D') &\geq \mathbf{E}_{v \leftarrow D} [(1 - \epsilon)p'_v] \\ &\geq (1 - \epsilon) \mathbf{E}_{v \leftarrow D} [p_v - \delta(\phi'_v)/\epsilon] \\ &\geq (1 - \epsilon) \text{REV}(D) - (1 - \epsilon)\bar{\delta}/\epsilon \end{aligned}$$

□

**Corollary 1.** *For a given partition of  $[n]$ ,  $S \sqcup T$ , and distribution  $D$  over subadditive valuations, define  $D_S$  to be  $D$  restricted to items in  $S$ , and  $D^+$  to first sample  $v \leftarrow D$ , and output  $v^+(\cdot)$  with  $v^+(U) = v(U \cap S) + v(U \cap T)$ . Then for all  $\epsilon \in (0, 1)$ ,  $\text{REV}(D) \leq \frac{\text{REV}(D^+)}{1 - \epsilon} + \frac{\text{VAL}(D_S)}{\epsilon}$ .*

*Proof.* By monotonicity,  $v(U) \geq v(U \cap T)$  for all  $U, T$ . Therefore,  $v'(U) - v(U) \leq v(U \cap S) \leq v(S)$  for all  $U$ . Furthermore, by subadditivity, we have  $v'(U) \geq v(U)$  for all  $U$ . Together, this means that  $D$  and  $D'$  are coupled so that we can set  $\delta(U) \leq v(S)$  for all  $U$ . Therefore, we may set  $\bar{\delta} = \text{VAL}(D_S)$  in the hypothesis of Lemma 4. The corollary follows by rearranging the inequality. □

**Lemma 5.** (“Approximate Core Decomposition”)

*For any distribution  $D$  that is subadditive over independent items, and any  $0 < \epsilon < 1$ ,*

$$\text{REV}(D) \leq \left( \frac{1}{\epsilon} + \frac{1}{1 - \epsilon} \right) \text{VAL}(D_\emptyset^C) + \frac{1}{1 - \epsilon} \sum_{A \subseteq [n]} p_A \text{REV}(D_A^T).$$

*Proof.* By the Approximate Marginal Mechanism Lemma (Lemma 2),

$$\text{REV}(D_A) \leq \left( \frac{1}{\epsilon} + \frac{1}{1 - \epsilon} \right) \text{VAL}(D_A^C) + \frac{1}{1 - \epsilon} \text{REV}(D_A^T)$$

Also, for any  $A \subseteq [n]$ ,

$$\text{VAL}(D_A^C) \leq \text{VAL}(D_\emptyset^C)$$

Finally, by sub-domain stitching (Lemma 1):

$$\begin{aligned} \text{REV}(D) &\leq \sum_{A \subseteq [n]} p_A \text{REV}(D_A) \\ &\leq \sum_{A \subseteq [n]} p_A \left( \left( \frac{1}{\epsilon} + \frac{1}{1 - \epsilon} \right) \text{VAL}(D_A^C) + \frac{1}{1 - \epsilon} \text{REV}(D_A^T) \right) \\ &\leq \left( \frac{1}{\epsilon} + \frac{1}{1 - \epsilon} \right) \text{VAL}(D_\emptyset^C) + \frac{1}{1 - \epsilon} \sum_{A \subseteq [n]} p_A \text{REV}(D_A^T) \end{aligned}$$

□

## C Omitted Proofs from Section 3.2

**Lemma 6.** *Let  $D$  be any distribution that is subadditive over independent items where each  $v(\{i\}) \in [0, c]$  with probability 1. Then  $D$  is  $c$ -Lipschitz.*

*Proof.* For any  $\vec{x}, \vec{y}, S, T$ , let  $U = \{i \in S \cap T \mid x_i = y_i\}$ . Because of no externalities, we must have  $V(\vec{x}, U) = V(\vec{y}, U)$ , which we will denote by  $B$ . By monotonicity, we must have  $V(\vec{x}, S), V(\vec{y}, T) \geq B$ . By subadditivity and the fact that each  $V(\vec{x}, \{i\}) \leq c$ , we have  $V(\vec{x}, S) \leq c(|S| - |U|) + B$ . Similarly, we have  $V(\vec{y}, T) \leq c(|T| - |U|) + B$ . It's also clear that  $|S| - |U| \leq |S \cup T| - |S \cap T| + |\{i \in S \cap T : x_i \neq y_i\}|$ , and that  $|T| - |U| \leq |S \cup T| - |S \cap T| + |\{i \in S \cap T : x_i \neq y_i\}|$ . So we also must have  $V(\vec{x}, S), V(\vec{y}, T) \leq B + c(|S \cup T| - |S \cap T| + |\{i \in S \cap T : x_i \neq y_i\}|)$ . Therefore  $V(\vec{x}, S), V(\vec{y}, T) \in [B, B + c(|S \cup T| - |S \cap T| + |\{i \in S \cap T : x_i \neq y_i\}|)]$ , completing the proof.  $\square$

**Lemma 7.**  $\text{SREV}_{\vec{p}}^* \geq t \cdot p_\emptyset (1 - p_\emptyset)$ . *In particular, if we choose  $t$  so that  $p_\emptyset = 1/2$ , then  $\text{SREV}_{\vec{p}}^* \geq t/4$ .*

*Proof.* Clearly  $\text{REV}_{p_i}(D_i) \geq p_i t$ , as we could set a price of  $t$  for item  $i$ . So  $\text{SREV}_{\vec{p}}^* = p_\emptyset \sum_i \text{REV}_{p_i}(D_i) \geq p_\emptyset t \sum_i p_i$ . Finally, we observe that  $\sum_i p_i$  is exactly the expected number of items in the tail, and  $p_\emptyset$  is the probability that zero items are in the tail. So we clearly have  $\sum_i p_i \geq 1 - p_\emptyset$ .  $\square$

## D Omitted Proofs from Section 3.3

**Lemma 8.**

$$\text{REV}(D) \leq 6n^{\log_2 6} \sum_i \text{REV}(D_i).$$

*Proof.* Babaioff et al. [1] prove that  $\text{REV} \leq n \sum_i \text{REV}(D_i)$  for an additive buyer by recursively reducing the number of items by one at each step. Unfortunately, each step of the induction uses the Marginal Mechanism Lemma; when applying the approximate variant for subadditive valuations, we would incur an exponential factor.

Instead, we use a slightly more complicated argument along the lines of Hart and Nisan [12] that halves the number of items in each step. Let  $S$  and  $T$  be a partition of  $[n]$  into subsets of size at most  $\lceil n/2 \rceil$ . Let  $D^{S \geq T}$  be the distribution over valuations which is the same as  $D$  whenever  $v(S) \geq v(T)$ , and has valuation zero otherwise. Similarly, let  $D^{S < T}$  be the distribution which is equal to  $D$  on  $v(S) < v(T)$ . Then by sub-domain stitching (Lemma 1) we have,

$$\text{REV}(D) \leq \text{REV}(D^{S \geq T}) + \text{REV}(D^{S < T}). \quad (4)$$

Now, by the Marginal Mechanism Lemma,

$$\text{REV}(D^{S \geq T}) \leq \left( \frac{1}{\epsilon} + \frac{1}{1 - \epsilon} \right) \text{VAL}(D_T^{S \geq T}) + \frac{1}{1 - \epsilon} \mathbf{E}_{v_T \leftarrow D_T^{S \geq T}} \left[ \text{REV}(D_S^{S \geq T} \mid v_T) \right] \quad (5)$$

One mechanism for selling items in  $S$  is to sample  $v_T \leftarrow D_T^{S \geq T}$ , and then use a mechanism that achieves  $\text{REV}(D_S^{S \geq T} \mid v_T)$ . Thus we have,

$$\text{REV}(D_S) \geq \max_{v_T} \text{REV}(D_S^{S \geq T} \mid v_T) \geq \mathbf{E}_{v_T \leftarrow D_T^{S \geq T}} \left[ \text{REV}(D_S^{S \geq T} \mid v_T) \right]. \quad (6)$$

Another way to sell items in  $S$  is to again sample  $v_T \leftarrow D_T^{S \geq T}$ , and offer the entire bundle for price  $v_T$ . Therefore we also have,

$$\text{REV}(D_S) \geq \mathbf{E}_{v \leftarrow D} \left[ v(T) \mid (v(S) \geq v(T)) \right] = \mathbf{E}_{v \leftarrow D^{S \geq T}} [v(T)] = \text{VAL}(D_T^{S \geq T}). \quad (7)$$

Combining equations (5)-(7), we have

$$\text{REV}(D^{S \geq T}) \leq \left( \frac{1}{\epsilon} + \frac{2}{1-\epsilon} \right) \text{REV}(D_S)$$

By symmetry, the same holds for  $\text{REV}(D^{S < T})$  and  $\text{REV}(D_T)$ . Therefore using (4),

$$\text{REV}(D) \leq \left( \frac{1}{\epsilon} + \frac{2}{1-\epsilon} \right) (\text{REV}(D_S) + \text{REV}(D_T))$$

Applying the recursion  $\lceil \log n \rceil$  times, we have

$$\text{REV}(D) \leq \left( \frac{1}{\epsilon} + \frac{2}{1-\epsilon} \right)^{\log_2 n + 1} \sum_i \text{REV}(D_i) = \left( \frac{1}{\epsilon} + \frac{2}{1-\epsilon} \right) n^{\log_2(\frac{1}{\epsilon} + \frac{2}{1-\epsilon})} \sum_i \text{REV}(D_i)$$

Choosing  $\epsilon = 1/2$  yields  $\left( \frac{1}{\epsilon} + \frac{2}{1-\epsilon} \right) = 6$ . □

**Proposition 2.** Recall that  $p_i = \Pr[v(\{i\}) > t]$ , and  $p_\emptyset = \prod_i (1 - p_i)$ . Then

$$\sum_A p_A \text{REV}(D_A^T) \leq \frac{6}{p_\emptyset} (1 + 7 \ln(1/p_\emptyset) + 6 \ln(1/p_\emptyset)^2 + \ln(1/p_\emptyset)^3) \cdot \text{SREV}_p^*$$

In particular, if we choose  $t$  so that  $p_\emptyset = 1/2$ , then  $\sum_A p_A \text{REV}(D_A^T) \leq 109 \cdot \text{SREV}_p^*$

*Proof.* Our proof builds on the intuition that the number of items in the tail is typically very small. By Lemma 8, we have that

$$\begin{aligned} \sum_{A \subseteq [n]} p_A \text{REV}(D_A^T) &\leq \sum_{A \subseteq [n]} p_A 6 |A|^{\log_2 6} \sum_{i \in A} \text{REV}(D_i^T) \\ &= 6 \sum_{i \in [n]} p_i \text{REV}(D_i^T) \sum_{A \ni i} |A|^{\log_2 6} p_A / p_i \end{aligned} \quad (8)$$

For any  $i$ , the expression  $\sum_{A \ni i} |A| p_A / p_i$  is also the expected number of items in the tail, conditioning on  $i$  being in the tail. Similarly,  $\sum_{A \ni i} |A|^{\log_2 6} p_A / p_i$  is the expected value of  $(\# \text{ items})^{\log_2 6}$ . Let  $b_j$  be the indicator random variable that is 1 whenever item  $j$  is in the tail. Noting that  $\log_2 6 < 3$  and each  $b_j$  is 1 with probability exactly  $p_j$  and the  $b_j$ 's are independent, we have:

$$\begin{aligned} \sum_{A \ni i} |A|^{\log_2 6} p_A / p_i &\leq \mathbb{E}_{b_j} \left[ \left( 1 + \sum_{j \neq i} b_j \right)^3 \right] \\ &= \mathbb{E}_{b_j} \left[ 1 + 3 \left( \sum_{j \neq i} b_j \right) + 3 \left( \sum_{j \neq i} b_j \right)^2 + \left( \sum_{j \neq i} b_j \right)^3 \right] \\ &= 1 + 3 \mathbb{E} \left[ \sum_{j \neq i} b_j \right] + 3 \mathbb{E} \left[ \sum_{j \neq i} b_j^2 + \sum_{k \neq j \neq i} b_j b_k \right] \\ &\quad + \mathbb{E} \left[ \sum_{j \neq i} b_j^3 + 3 \sum_{k \neq j \neq i} b_j^2 b_k + \sum_{l \neq k \neq j \neq i} b_j b_k b_l \right] \\ &= 1 + 7 \mathbb{E} \left[ \sum_{j \neq i} b_j \right] + 6 \mathbb{E} \left[ \sum_{k \neq j \neq i} b_j b_k \right] + \mathbb{E} \left[ \sum_{l \neq k \neq j \neq i} b_j b_k b_l \right] \end{aligned} \quad (9)$$

$$\leq 1 + 7 \sum_{j \neq i} p_j + 6 \left( \sum_{j \neq i} p_j \right)^2 + \left( \sum_{j \neq i} p_j \right)^3 \quad (10)$$

(9) follows because  $b_j \in \{0, 1\}$ . We continue to bound the last line as a function of just  $p_\emptyset$ . Note that  $e^{-\sum_i p_i} \geq \prod_i (1 - p_i) = p_\emptyset$ , and therefore we have  $\sum_i p_i \leq \ln(1/p_\emptyset)$ . Combining with (8) and (10), we have:

$$\sum_{A \subseteq [n]} p_A \text{REV}(D_A^T) \leq 6(1 + 7 \ln(1/p_\emptyset) + 6 \ln(1/p_\emptyset)^2 + \ln(1/p_\emptyset)^3) \cdot \sum_{i \in [n]} p_i \text{REV}(D_i^T)$$

Now, we have to interpret  $p_i \text{REV}(D_i^T)$ . We claim that in fact this is exactly  $\text{REV}_{p_i}(D_i)$ . Why? It's clear that the optimal reserve for  $D_i^T$  is at least  $t$ , as  $D_i^T$  is not supported below  $t$ . It's also easy to see that for any reserve  $r_i \geq t$ , that the revenue obtained by selling to  $D_i^T$  is exactly  $\Pr[v(\{i\}) > r_i]/p_i$ , and therefore the same  $r_i \geq t$  that is optimal for  $D_i$  is also optimal for  $D_i^T$ , and  $p_i \text{REV}(D_i^T) = \text{REV}_{p_i}(D_i)$ . Therefore,

$$\sum_{A \subseteq [n]} p_A \text{REV}(D_A^T) \leq 6(1 + 7 \ln(1/p_\emptyset) + 6 \ln(1/p_\emptyset)^2 + \ln(1/p_\emptyset)^3) \cdot \sum_{i \in [n]} \text{REV}_{p_i}(D_i)$$

Plug in  $\text{SREV}_{\bar{p}}^* = p_\emptyset \sum_{i \in [n]} \text{REV}_{p_i}(D_i)$  to complete the proof.  $\square$

## E Approximate Marginal Mechanism and Core Decomposition for Many Bidders

In this section we provide a complete proof of the Marginal Mechanism and Core Decomposition lemmas that apply to many bidders. The statements are essentially the same, but require more extensive notation which we develop in the next subsection.

The main technical challenge is extending Lemma 4 to many buyers; recall that this is the lemma that lower bounds the revenue from a distribution  $D^+$  of additive-across-a-partition valuations, in terms of the revenue from the original distribution  $D$  and the expected difference  $\bar{\delta}$  between  $D$  and  $D^+$ . I.e. given a black-box mechanism for  $D$ , we would like to create a new mechanism for  $D^+$ . How can we preserve incentive compatibility when the buyers have different incentives? In Appendix B we do this for a single buyer by giving the buyer the outcome (allocation and price), of the possible outcomes for all types in the original mechanism, which is optimal for her new valuation. Incentive compatibility is now guaranteed, and we simply need to bound the revenue.

When multiple buyers are involved there is a problem with this approach: incentive compatibility of any buyer depends also on the distribution of types reported by other buyers; thus we cannot simply let any buyer report any type she wants. To overcome this challenge, we use a technique due to [3, 15, 11] which guarantees that each buyer observes the correct distribution of types on the other buyers. For buyer  $i$ , we sample an additional  $r - 1$  replica types from  $D_i^+$ , and  $r$  surrogate types from  $D_i$ . Given the real buyer's type, we create (in an incentive compatible manner) a complete matching between replicas and surrogates. Since the real buyer's type is sampled from the same distribution as the replicas, she is equally likely to be matched to any of the surrogates. Thus the distribution over the surrogate type that is matched to the buyer  $i$  is exactly the original distribution  $D_i$ ; i.e. all other buyers observe the "correct" distribution. The mechanism is now Bayesian incentive compatible, and further analysis shows that the lower bound on the revenue is preserved.

### E.1 Notation and statement

There are  $m$  bidders and  $n$  items. We now say that an item  $i$  is in the tail if there is *any* buyer who values item  $i$  above the core-tail cutoff, and item  $i$  is in the core if every buyer values it below the cutoff.  $D$  is now the joint distribution of all buyers valuation functions. We study Bayesian Incentive Compatible (BIC) mechanisms. A mechanism is BIC if it is in every buyer's

best interest to report truthfully, conditioned on other buyers reporting truthfully as well. Formally, let  $\phi_j(v)$  denotes the (random) allocation awarded to bidder  $j$  when reporting type  $v$ ; we slightly abuse notation by letting  $v(\phi)$  denote the expected utility, over any randomness in the mechanism and the other bidders sampling their types, that a bidder with type  $v$  gains from a random allocation  $\phi$ ; let  $p_j(v)$  denote the expected price paid by bidder  $j$  when reporting type  $v$  over the same randomness. Finally, a mechanism is BIC iff for all  $j, v, v'$ ,  $v(\phi_j(v)) - p_j(v) \geq v(\phi_j(v')) - p_j(v')$ . We use the following notation.

- $D_j$ : The distribution of  $v_j(\cdot)$ , the valuation function for bidder  $j$ . We assume that  $D$  is a product distribution. That is,  $D = \times_j D_j$ .
- $D_A$ : the distribution  $D$ , conditioned on  $A$  being exactly the set of items in the tail.
- $D_A^T$ : the distribution  $D_A$  restricted just to items in the tail (i.e.  $A$ ).
- $D_A^C$ : the distribution  $D_A$  restricted just to items in the core (i.e.  $\bar{A}$ ).
- $p_i$ : the probability that element  $i$  is in the tail.
- $p_A$ : the probability that  $A$  is exactly the set of items in the tail (note that  $p_{\{i\}} \neq p_i$ ).
- $\text{REV}(D)$ : The maximum revenue obtainable via a BIC mechanism from buyers with valuation profile drawn from  $D$ .
- $\text{VAL}(D) = \mathbb{E}_{v \leftarrow D} \left[ \max_{S_1 \sqcup \dots \sqcup S_m} \{ \sum_j v_j(S_j) \} \right]$ : the expected welfare of the VCG mechanism when buyers are drawn from  $D$ .
- $\text{REV}^M(D)$ : The revenue of a BIC mechanism  $M$  when buyers drawn from  $D$  play truthfully.

We are finally ready to state the approximate core decomposition lemma for many buyers:

**Lemma 9.** *For any distribution  $D = \times_j D_j$  with each  $D_j$  subadditive over independent items, and any  $0 < \epsilon < 1$ ,*

$$\text{REV}(D) \leq \left( \frac{1}{\epsilon} + \frac{1}{1-\epsilon} \right) \text{VAL}(D_\emptyset^C) + \frac{1}{1-\epsilon} \sum_{S \subseteq [n]} p_A \text{REV}(D_A^T)$$

*Proof.* Follows from Theorem 4 below, together with the arguments used in Section 3.1 and in [1].  $\square$

## E.2 The mechanism

First, we describe the reduction we will use (which is essentially the same as the  $\epsilon$ -BIC to BIC reduction used in [11], but without some technical hardships since we aren't concerned with runtime - our proof never actually runs this procedure). Below,  $D^+$  denotes any product distribution that first-order stochastically dominates  $D$ , and  $\delta_j(\cdot)$  denotes the random function  $v_j^+(\cdot) - v_j(\cdot)$  when couples  $v^+$  and  $v$  are sampled jointly from  $D^+$  and  $D$ . Note that  $\delta_j(S) \geq 0$  for all  $\delta_j, S$ . We will also abuse notation and refer to  $\delta_j$  as the distribution over  $\delta_j(\cdot)$  as well (so we can write terms like  $\text{VAL}(\delta)$ ).

### Phase 1, Surrogate Sale:

1. Let  $M$  be any *BIC* mechanism for buyers from  $D$ . Multiply all prices charged by  $M$  by  $(1 - \epsilon)$  and call the new mechanism  $M^\epsilon$ . Interpret the  $\epsilon$  fraction of prices given back as rebates.
2. For each bidder  $j$ , create  $r - 1$  replicas sampled i.i.d. from  $D_j^+$  and  $r$  surrogates sampled i.i.d. from  $D_j$ . Let  $r \rightarrow \infty$ .
3. Ask each bidder to report  $v_j(\cdot)$ .
4. Create a weighted bipartite graph with replicas (and bidder  $j$ ) on the left and surrogates on the right. The weight of an edge between a replica (or bidder  $j$ ) with type  $r_j(\cdot)$  and surrogate of type  $s_j(\cdot)$ , is the utility of  $r_j$  for the expected outcome of  $M^\epsilon$  when reporting  $s_j$ . That is, the weight of the edge is  $r_j(\phi_j^\epsilon(s_j)) - p_j^\epsilon(s_j)$ .
5. Compute a maximum matching and VCG prices in this bipartite graph; we henceforth refer to it as the *VCG matching*. If a replica (or bidder  $j$ ) is unmatched in the VCG matching, add an edge to a random unmatched surrogate. (Notice that some replicas may indeed be unmatched if the gain negative utility from the allocation and prices corresponding to some surrogates.) The surrogate selected for bidder  $j$  is whomever she is matched to.

### Phase 2, Surrogate Competition:

1. Let  $s_j$  denote the surrogate chosen to represent bidder  $j$  in phase one, and let  $\vec{s}$  denote the entire surrogate profile (i.e. the ones matched to the real buyers). Have the surrogates play  $M^\epsilon$ .
2. If bidder  $j$  was matched to her surrogate through VCG, charge them the VCG price and award them  $M_j^\epsilon(\vec{s})$ . Recall that this has an allocation and price component; the price is added onto the VCG price. If bidder  $j$  was matched to a random surrogate after VCG, award them nothing and charge them nothing.

**Theorem 4.** *Let  $M'$  denote the mechanism of the process above, starting from any BIC mechanism  $M$  for consumers drawn from  $D$ . Then  $M'$  is BIC for consumers drawn from  $D^+$ . Furthermore, for any desired  $\epsilon \in (0, 1)$ , we can have:  $\text{REV}^{M'}(D^+) \geq (1 - \epsilon) \cdot (\text{REV}^M(D) - \text{VAL}(\delta)) / \epsilon$ .*

In the theorem statement above, note that by  $\text{VAL}(\delta)$ , we mean the expected welfare of the VCG mechanism when buyers with types distributed according to  $\delta = \times_j \delta_j$  play.

### E.3 Proof outline

The reduction is nearly identical to the reduction employed in [11] (which is itself inspired by [3, 15]). We provide complete proofs of all claims for completeness, noting that many of these claims can also be found in [3, 15, 11]. We provide appropriate citations by each statement. Below is the proof outline, taken from [11].

1. If bidder  $j$  plays  $M'$  truthfully, then the distribution of surrogates matched to bidder  $j$  is  $D_j$ . Therefore, the value of each edge is calculated correctly as the expected utility of a replica with type  $r$  for being represented by a surrogate of type  $s$  in  $M^\epsilon$ .
2. Because each bidder is participating in a VCG auction for their surrogate, and the value of each edge is calculated correctly, whenever all other bidders tell the truth, it is in bidder  $j$ 's interest to tell the truth as well. Therefore,  $M'$  is BIC.
3. The revenue made from bidder  $j$  is at least the price paid by their surrogate if bidder  $j$  is matched in VCG, and 0 otherwise.

4. There exists a near-perfect matching that matches each replica to a nearly-identical surrogate (modulo  $\delta_j(\cdot)$ ). If VCG used this matching, we would have  $\text{REV}^{M'}(D) = (1 - \epsilon)\text{REV}^M(D)$ .
5. The rebates allow us to bound the revenue lost by selecting the VCG matching instead of this near-perfect matching as a function of  $\text{VAL}(\delta)$  and  $\epsilon$ .

We proceed to state the formal claims associated with steps 1 through 5.

### Step 1: The surrogate distributions

**Lemma 10.** ([15]) *If bidder  $j$  plays  $M'$  truthfully, then the distribution of the surrogate selected for bidder  $j$  is exactly  $D_j$ .*

*Proof.* Imagine sampling replicas and surrogates for bidder  $j$  in the following way instead. First, sample the  $r$  types for the left-hand side of the bipartite graph i.i.d. from  $D_j^+$  and the  $r$  types for the right-hand side i.i.d. from  $D_j$ . Then, find the max-weight matching between types, completing it by randomly adding edges from unmatched nodes on the left to unmatched nodes on the right to form a perfect matching. Then, declare all of the  $r$  right-hand types surrogates, randomly select one of the left-hand types to be bidder  $j$ , and declare the remaining  $r - 1$  as replicas. Note that sampling in this order yields the correct distribution of replicas, surrogates, and bidder  $j$ , as long as bidder  $j$  reports truthfully. Furthermore, it is clear now that the distribution of the surrogate selected for bidder  $j$  after sampling in this order is exactly  $D_j$ : once the matching is fixed, we simply pick a random left-hand type and output its partner. So essentially we are drawing  $r$  i.i.d. samples from  $D_j$  and selecting one at random. Clearly this is the same distribution as  $D_j$ .  $\square$

### Step 2: Bayesian incentive compatibility

**Corollary 8.** ([15])  *$M'$  is BIC.*

*Proof.* Fix any bidder  $j$  and assume all others report truthfully. By Lemma 10, the distribution of all other surrogates matches  $D_{-j}$  exactly, so the weight of the edge between each replica (and bidder  $j$ ) and each surrogate correctly computes the value of that replica for being represented by that surrogate in  $M$ . As bidder  $j$  is just participating in a truthful VCG mechanism against the replicas for the surrogates, and all values are computed correctly (conditioned on other buyers telling the truth),  $M'$  is BIC.  $\square$

### Step 3: A good matching implies high revenue

**Proposition 4.** ([11]) *Conditioning on right-hand types (surrogates)  $\{s_k\}_{k \in [r]}$  and left-hand types (replicas plus bidder  $j$ )  $\{r_k\}_{k \in [r]}$ , the expected payment of bidder  $j$  is at least*

$$\sum_{k \text{ s.t. } s_k \text{ is matched in VCG}} p_j^\epsilon(s_k)/r.$$

*Proof.* Conditioned on the left and right-hand types being sampled, but having not yet decided which left-hand type is bidder  $j$ , the surrogate matched to bidder  $j$  is just a random surrogate. So each surrogate  $s_k$  is matched to bidder  $j$  with probability  $1/r$ . Furthermore, bidder  $j$  will pay the price  $p_j^\epsilon(s_k)$  whenever his matched surrogate was selected by VCG (and not the random edges afterwards). Therefore, bidder  $j$  pays at least  $\sum_{\{k|s_k \text{ is matched in VCG}\}} p_j^\epsilon(s_k)/r$ . The reason this is not tight is because it does not count the additional payments of the VCG mechanism.  $\square$

#### Step 4: Existence of a near-perfect matching

Our next goal is to show that there exists a near-perfect matching that only matches replicas and surrogates that are “close.” For any  $\gamma > 0$  and two types  $v$  and  $v'$  drawn from  $D_j$ , we say that  $v$  and  $v'$  are  $\gamma$ -equivalent if for all  $S \subseteq [n]$ , there exists an integer  $z(S)$  such that  $\{v(S), v'(S)\} \subseteq [(1 + \gamma)^{z(S)}, (1 + \gamma)^{z(S)+1}]$ . It’s easy to see that this defines an equivalence relation. For a type  $r_j$  drawn from  $D_j^+$ , let  $r'_j$  denote its couple drawn from  $D_j$ . We say that two types drawn from  $D_j^+$  are  $\gamma$ -equivalent if their couples are  $\gamma$ -equivalent, and that  $r_j$  drawn from  $D_j^+$  is  $\gamma$ -equivalent to  $s_j$  drawn from  $D_j$  if  $r'_j$  and  $s_j$  are  $\gamma$ -equivalent (basically we are putting replicas in equivalence classes based on their couples).

The following lemma from [15] bounds the number of unmatched surrogates:

**Lemma 11.** (Lemma 3.7 in [15]) *The expected number of unmatched surrogates in a maximal matching that only matches equivalent replicas and surrogates when types are split into at most  $\beta$  possible equivalence classes is at most  $O(\sqrt{\beta r})$ .*

In the next lemma we use Lemma 11 to lower-bound the revenue obtained from the matched replicas.

**Lemma 12.** (Implicit in [15]) *For any  $\gamma > 0$ , let  $W$  denote any maximal matching of replicas to surrogates (for all bidders) that only matches  $\gamma$ -equivalent types. Then as  $r \rightarrow \infty$ , we get  $E[\sum_j \sum_{\{k|s_{jk} \text{ is matched in } W\}} p_j^\epsilon(s_{jk})] \geq (1 - \gamma)\text{REV}(M^\epsilon)$ .*

*Proof.* At a high level, the proof is straightforward: for a fixed equivalence class, the distribution of the number of replicas and surrogates in that class is the same. So as we take more and more i.i.d. samples, the number of each concentrates very tightly around its expectation, so not many types are unmatched. Showing this formally is somewhat technical.

Let’s focus on a specific equivalence class  $C$  for a fixed bidder  $j$ . There is some probability  $q_C$  that a type drawn from  $D_j$  lands in class  $C$ . Let  $\text{REV}^q(M^\epsilon)$  denote the revenue obtained by  $M^\epsilon$  only counting payments by types in an equivalence class  $C$  such that  $q_C \geq q$  and  $v([n]) \leq 1/q$  for all  $v \in C$ . It’s clear that  $\lim_{q \rightarrow 0} \text{REV}^q(M^\epsilon) = \text{REV}(M^\epsilon)$ , as the revenue obtained from equivalence classes with  $q_C = 0$  is exactly 0 and the revenue obtained from all other equivalence classes is eventually counted for sufficiently small  $q$ . So we can pick a  $q > 0$  such that  $\text{REV}^q(M^\epsilon) \geq (1 - \gamma/2)\text{REV}(M^\epsilon)$ . Note now that there can only be finitely many equivalence classes counted towards  $\text{REV}^q(M^\epsilon)$  (in particular, at most  $1/q$  per bidder), and that the maximum payment by a type in any such equivalence class is at most  $1/q$  (by individual rationality).

So now the probability that a surrogate or replica is sampled to be in a counted equivalence class is at least  $q$  (there must be at least one such equivalence class) and there are at most  $1/q$  equivalence classes. So we may apply Lemma 11 to see that for a single bidder, the expected number of unmatched surrogates from equivalence classes that count is at most  $O(\sqrt{r/q})$ . As each such surrogate pays at most  $1/q$ , the total revenue lost in expectation from unmatched surrogates in equivalence classes that count is at most  $O(1/\sqrt{rq^3})$  (due to Proposition 4). Summing up across all bidders, the revenue lost is at most  $O(m/\sqrt{rq^3})$ . In addition, the total revenue lost in expectation from unmatched surrogates across all bidders from equivalence classes that didn’t count is clearly at most  $(\gamma/2)\text{REV}(M^\epsilon)$  by choice of  $q$ . So the total revenue lost from unmatched surrogates in this matching is at most  $(\gamma/2)\text{REV}(M^\epsilon) + O(m/\sqrt{rq^3})$ . As  $r \rightarrow \infty$ , the second term approaches zero, completing the proof.  $\square$

#### Step 5: The VCG matching is almost as good

Combining Proposition 4 and Lemma 12 says the following: if only this nearly-perfect matching was the one selected by VCG, then we would know that  $\text{REV}^{M'}(D)$  was good. But for all we

know, VCG may choose to leave many surrogates unmatched if it helps improve the welfare of the replicas. The key is to show that not many surrogates can be unmatched, due to the rebates.

Let  $V$  denote the VCG matching and  $W$  denote the matching guaranteed by Lemma 12. Then there is a disjoint set of augmenting paths and cycles that transform  $W$  into  $V$ . As  $V$  is a max-weight matching, *all of these augmenting paths and cycles must have non-negative weight*. It is easy to see that augmenting cycles do not change the set of matched surrogates, and therefore do not affect the revenue. We therefore want to study augmenting paths that unmatch a surrogate.

If an augmenting path unmatches  $s_{jk}$ , then no replica is receiving the rebate awarded to  $s_{jk}$  any more. Because VCG is choosing the max-weight matching, it must be the case that the benefit of switching every other edge along the path outweighs the cost of losing the rebate awarded to  $s_{jk}$ . This is where we make use of the fact that each replica is matched to a surrogate that is nearly identical to them, except for an additive  $\delta_{jk}(\cdot)$ . Because  $M$  is BIC, we can bound the expected gain of switching a replica who is matched to a nearly identical surrogate to any other surrogate using  $\delta(\cdot)$ . Therefore, each surrogate that gets unmatched by an augmenting path “claims” many replicas to be in its augmenting path. The argument shows that in fact it takes several replicas to make a positive weight augmenting path, and therefore not many surrogates can be unmatched.

**Lemma 13.** (*Ideas from [11]*) *If  $U_j$  denotes the set of surrogates that are matched in  $W$  but not  $V$ , and  $T_j$  denotes the set of surrogates matched in  $V$  but not  $W$ , (for bidder  $j$ ), then  $\mathbb{E}[\frac{1}{r} \sum_j \sum_{s \in U_j} p_j(s) - \sum_{s \in T_j} p_j(s)] \leq \text{VAL}(\delta)/\epsilon$ .*

*Proof.* Consider any augmenting path that unmatches a surrogate  $s$  in  $W$  and (possibly) matches a new surrogate  $s'$ . For simplicity of notation, if no new surrogate is matched, we let  $s'$  denote a null type that receives no items and pays no price to  $M$ . We break the contribution of edges in this path into two parts, the first coming just from the rebates awarded to the surrogates and the second coming from the allocation and original price paid. It is easy to see that the contribution of rebates to the weight of the augmenting path is exactly  $\epsilon p_j(s) - \epsilon p_j(s')$ . Now we analyze the weight of the path coming from the second part. We can compute the weight by summing over all replicas  $r_j$  in the path of their utility for their new surrogate minus their utility for their old surrogate. Note that any augmenting path that unmatches a surrogate cannot possibly add an edge to a replica who was unmatched in  $W$ . Since  $M$  is BIC, for any replica  $r_j$  that was matched to  $s_j$  in  $W$  and moved to  $s'_j$  in  $V$ , we have:

$$s_j(\phi_j(s_j)) - p_j(s_j) \geq s_j(\phi_j(s'_j)) - p_j(s_j)$$

Using the fact that  $r_j$  and  $s_j$  are  $\gamma$ -equivalent, we get:

$$(1 + \gamma)r_j(\phi_j(s_j)) - p_j(s_j) \geq (1 - \gamma)r_j(\phi_j(s'_j)) - \delta_j(\phi_j(s'_j)) - p_j(s_j)$$

And rearranging terms yields:

$$r_j(\phi_j(s'_j)) - p_j(s_j) - r_j(\phi_j(s_j)) + p_j(s_j) \leq \gamma r_j(\phi_j(s_j)) + \gamma r_j(\phi_j(s'_j)) + \delta_j(\phi_j(s'_j))$$

Note that the left-hand side is exactly the increase in utility as we move  $r_j$  from  $s_j$  to  $s'_j$ . So now we know that the total increase in utility from moving all replicas across all bidders must outweigh the total decrease caused by unmatching surrogates. We therefore get:

$$\mathbb{E}\left[\frac{1}{r} \sum_j \sum_{s \in U_j} p_j(s) - \sum_{s \in T_j} p_j(s)\right] \leq \mathbb{E}\left[\frac{1}{r} \sum_j \sum_{r_j} \gamma r_j(\phi_j(s_j)) + \gamma r_j(\phi_j(s'_j)) + \delta_j(\phi_j(s'_j))\right]$$

Consider now the following allocation algorithm (*not* a mechanism, and certainly not truthful). When bidder  $j$  reports  $r_j$ , the algorithm selects the allocation output by  $M$  on input  $(s_1, \dots, s_m)$ . This is clearly a feasible allocation algorithm, and it's also clear that the term  $\mathbb{E}[\frac{1}{r} \sum_j \sum_{r_j} r_j(\phi_j(s_j))]$  exactly computes the expected welfare achieved by this algorithm when the type of buyer  $j$  is drawn from  $D_j$ . We can similarly define an allocation algorithm that replaces  $r_j$  with  $s'_j$ , and one that samples  $\delta_j \leftarrow D_j^+ - D_j$  and replaces  $\delta_j$  with  $s'_j$ . Now that we have concrete allocation algorithms that match the terms on the right-hand side exactly, we can bound it as:

$$\mathbb{E}[\frac{1}{r} \sum_j \sum_{r_j} \gamma r_j(\phi_j(s_j)) + \gamma r_j(\phi_j(s'_j)) + \delta_j(\phi_j(s'_j))] \leq 2\gamma \text{VAL}(D) + \text{VAL}(\delta)$$

Lastly, because this claim holds for all  $\gamma > 0$ , we may let  $\gamma \rightarrow 0$  and the right-hand bound becomes just  $\text{VAL}(\delta)$ . Some care must be taken if  $\text{VAL}(D) = \infty$ , but similar tricks to those used in the proof of Lemma 12 suffice. Essentially, one can take an increasing limit of truncations of  $D$ , calling them  $D^C$  (the same distribution  $D$  but replacing  $v(\cdot)$  with the 0-function if  $v([n]) > C$ ). Clearly all  $\text{VAL}(D^C)$  is finite for all  $C$ , and clearly  $\lim_{C \rightarrow \infty} \text{REV}^M(D^C) = \text{REV}^M(D)$  for any mechanism  $M$ . So one can use the above analysis on the bounded distributions  $D^C$  and take a limit.  $\square$

With the lemma above, our proof is complete. There is a high-cardinality matching  $W$  that provides revenue exactly  $\text{REV}^M(D)$ , if only it were chosen by VCG. But VCG may choose a different matching, and we may lose some revenue. The lemma bounds how much revenue can be lost, and provides the bound in the theorem.