Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Simple and Approximately Optimal Mechanism for a Buyer with Complements

Alon Eden, Michal Feldman, Ophir Friedler, Inbal Talgam-Cohen, S. Matthew Weinberg

informs
http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 69, No. 1, January–February 2021, pp. 188–206
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Crosscutting Areas

A Simple and Approximately Optimal Mechanism for a Buyer with Complements

Alon Eden,[a,b] Michal Feldman,[c,d] Ophir Friedler,[c] Inbal Talgam-Cohen,[e] S. Matthew Weinberg[b,f]

[a] Computer Science Department, Harvard University, Cambridge, Massachusetts 02139; [b] Simons Institute for the Theory of Computing, Berkeley, California 94720; [c] Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv 6997801, Israel; [d] Microsoft Research, Herzliya 4672513, Israel; [e] Computer Science Department, Technion, Haifa 3200003, Israel; [f] Computer Science Department, Princeton University, Princeton, New Jersey 08540

Contact: alonarden@gmail.com (AE); michal.feldman@cs.tau.ac.il (MF); ophirfriedler@gmail.com, https://orcid.org/0000-0002-7651-8223 (OF); italgam@cs.technion.ac.il, https://orcid.org/0000-0002-2838-3264 (IT-C); smweinberg@princeton.edu (SMW)

Received: May 31, 2018
Accepted: April 1, 2020
Published Online in Articles in Advance: December 15, 2020

Subject Classifications: Primary: games/group decisions: bidding/auctions; secondary: analysis of algorithms: suboptimal algorithms
Area of Review: Revenue Management and Market Analytics

https://doi.org/10.1287/opre.2020.2039

Abstract. We consider a revenue-maximizing seller with m heterogeneous items and a single buyer whose valuation for the items may exhibit both substitutes and complements. We show that the better of selling the items separately and bundling them together—guarantees a $\Theta(d)$-fraction of the optimal revenue, where d is a measure of the degree of complementarity; it extends prior work showing that the same simple mechanism achieves a constant-factor approximation when buyer valuations are subadditive (the most general class of complement-free valuations). Our proof is enabled by a recent duality framework, which we use to obtain a bound on the optimal revenue in the generalized setting. Our technical contributions are domain specific to handle the intricacies of settings with complements. One key modeling contribution is a tractable notion of "degree of complementarity" that admits meaningful results and insights—we demonstrate that previous definitions fall short in this regard.

Funding: This work was partially supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program [Grant 866132] (to M. Feldman), the Israel Science Foundation [Grant 336/18] (to I. Talgam-Cohen), and the Taub Family Foundation (to I. Talgam-Cohen). This work was supported by H2020 Marie Skłodowska-Curie Actions [Grant 708935], the National Science Foundation [Grant CCF 1717899] (to S. M. Weinberg), and the Israel Science Foundation [Grant 317/17] (to M. Feldman).

Keywords: mechanism design • revenue • approximation • complements

1. Introduction

Consider a revenue-maximizing seller with $m \geq 1$ items to sell to a single buyer. When there is just a single item, and the buyer's value is drawn from some distribution with Cumulative Distribution Function (CDF) F, seminal works of Myerson (1981) and Riley and Zeckhauser (1983) prove that the optimal mechanism is to simply set whatever price maximizes $p \cdot (1 - F(p))$. It is now well understood that beyond the single-item setting, the optimal mechanism suffers many undesirable properties that make it impractical, including randomization, nonmonotonicity, and others (Rochet and Chone 1998; Thanassoulis 2004; Pavlov 2011; Daskalakis et al. 2013, 2014; Hart and Nisan 2013; Briest et al. 2015; Hart and Reny 2015). Following the seminal work of Chawla et al. (2007), there is now a sizable body of research proving that the simple mechanisms we see in practice are in fact approximately optimal in quite general settings, helping to explain their widespread use (Chawla et al. 2010, 2015; Li and Yao 2013; Babaioff et al. 2014; Kleinberg and Weinberg 2014; Bateni et al. 2015; Rubinstein and Weinberg 2015; Yao 2015; Chawla and Miller 2016; Cai and Zhao 2017; Hart and Nisan 2017).

Still, prior work has largely been limited to additive or unit-demand buyers (a buyer's valuation is additive if $v(S) = \sum_{i \in S} v(\{i\})$; it is unit demand if $v(S) = \max_{i \in S}\{v(\{i\})\}$). Only recently have researchers begun tackling more complex valuation functions (Rubinstein and Weinberg 2015, Chawla and Miller 2016, Cai and Zhao 2017). Even these works have remained restricted to subclasses of *subadditive* valuations, also called *complement free* (a valuation is subadditive if $v(S \cup T) \leq v(S) + v(T)$ for all S, T). Although subadditive valuations are quite general, they can only capture interaction between items as *substitutes*. For example, if the items are pieces of furniture, a buyer's marginal value for a chair might decrease as her living space gets more crowded by other items. To date, no results in this line of work have modeled interactions between items as *complements*. For example, a buyer's value for a kitchen table might instead increase if she already has a chair. The goal of this paper is to study simple and approximately optimal mechanisms in domains where buyer valuations exhibit both substitutes and complements.

1.1. Running Examples

We now present two running examples, which will help motivate and exposit our model.

Example 1. Consider a college that sells courses (indexed by $[m] = \{1, 2, \ldots, m\}$) "a la carte" (i.e., courses are items for sale). The college also specifies diplomas that it awards (indexed by $[k]$), and for each diploma j, it specifies a set of courses S_j that a student must pass in order to receive j. Students may purchase courses as they like, but in order to receive a diploma, they must complete the set of courses required by that diploma. Denote by v_j the value of diploma j. The value of a student for a set of courses C is the total value of diplomas it covers (i.e., $\sum_{j:S_j \subseteq C} v_j$). The items (courses) exhibit complementarities because value is derived from diplomas containing multiple courses. At the same time, certain sets of courses may conflict, require prerequisites, or impose too much coursework. Therefore, the items also behave as substitutes (e.g., a student cannot successfully take two conflicting courses).

Example 2. Consider a recruiter who matches professionals to firms. The recruiter focuses on identifying and matching certain skills indexed by $[m]$ (say software developer, algorithms developer, data scientist, computer vision expert, quality assurance, etc.) to high-tech firms. A firm may be interested in maintaining a set of projects indexed by $[k]$, and each project $j \in [k]$ requires a set S_j of professionals to maintain. If v_j denotes the profit the firm will enjoy by completing project j, then the value of a firm for a set of experts C is the total value of projects they can cover (i.e., $\sum_{j:S_j \subseteq C} v_j$). As in Example 1, the items (experts) exhibit complementarities because value is derived from projects requiring multiple experts. Also as in Example 1, the items behave as substitutes—firms typically have limited hiring budget and/or physical space and cannot derive utility from recommended experts beyond this cap.

In both examples, observe that complementarities arise in a similar manner: Each individual item is a building block for an object that provides value. Substitutability arises in a similar manner as well: Certain subsets of items may conflict and be infeasible to simultaneously utilize. Because we view one of our main contributions as the formalization of an improved degree of complementarity notion, we dedicate the subsequent subsection to introducing this (and related) concepts.

1.2. How to Measure the Degree of Complementarity

Even for the traditionally simpler domain of *welfare* maximization, mechanisms for buyers with complements have only recently emerged (Abraham et al. 2012; Feige et al. 2015; Feldman et al. 2015, 2016). The main difficulty is that strong lower bounds are known for general valuations (Lehmann et al. 2002, Nisan and Segal 2006), so the precise degree to which buyer valuations exhibit substitutes or complements must be explicitly modeled in order to achieve tractability. Interestingly, strong positive results are possible in the complete absence of complements and no restriction on the degree of substitutability (Dobzinski 2007, Feige 2009, Dobzinski et al. 2010, Feldman et al. 2013, Devanur et al. 2015) but not vice versa: Many strong lower bounds still exist in the absence of substitutes but with arbitrary complementarity (Lehmann et al. 2002, Abraham et al. 2012, Morgenstern 2015, Feldman et al. 2016).

Although arbitrary substitutability may apply in quite general settings, valuations with arbitrary complements often misrepresent the settings at hand. In settings that involve complements (e.g., shoes, airport arrival and departure slots, radio spectrum allocation, labor markets with heterogeneous skills), it is rarely the case that *all* elements complement one another. Interestingly, the right definition of "degree of complementarity" differs between environments. The goal is thus to define valuation classes that allow reasoning about complements in a meaningful and sufficiently general way.

1.2.1. Building Block: The Positive Hypergraph Model.

Our proposed model of complementarity begins with the previously studied positive hypergraph (PH) model. A PH valuation $v(\cdot)$ is such that there exists another set function $w(\cdot)$ so that $v(S) = \sum_{T \subseteq S} w(T)$, and $w(T) \geq 0$ for all T. Note that such valuations exhibit complementarities but not substitutes. A PH valuation has PH degree d if for all $|T| > d$, $w(T) = 0$ (i.e., a valuation incorporates more complementarity behavior if larger sets of items have synergy). A prominent example for such a source of complementarity is the Federal Communications Commission (FCC) spectrum auctions, where purchasing licenses for the same band of spectrum in adjacent geographical regions increases the overall value of the purchase. Each region can be associated with an element, and each subset of geographically adjacent regions T may be associated with $w(T)$, the complementary benefit from owning the band at exactly the region set T. In Abraham et al. (2012), the authors provided algorithms and mechanisms for social welfare maximization with guarantees that degrade gracefully with d. It is worth noting that even for PH valuations of degree 2, it may be that *every pair of items* is complement. Note that in both Examples 1 and 2, the valuation function of the buyer is a PH valuation defined by $w(S_j) = v_j$ for all j, and $w(T) = 0$ otherwise. Therefore, the PH degree of complementarity would be $\max_j |S_j|$.

1.2.2. Shortcomings of the PH Model (and Alternatives). In the PH model, the *number of items* plays a crucial role in the degree of complementarity. For example, a buyer that is interested only in the grand bundle of all items (and no proper subset) has the highest degree of complementarity (in the context of Example 1, consider a college that sells many courses but certifies only a single diploma—both previous measures label the corresponding valuation function with the highest possible degree of complementarity of m). However, when selling to such a buyer, one can treat the grand bundle as a single item, and thus, no complementarity issues arise (so the "intrinsic" degree of complementarity is low). That is, previous measures of complementarity deem this setting the *most* complex, whereas from a revenue-maximizing perspective, it is actually the *least* complex. This observation suggests that a different perspective on complementarity is needed for revenue-maximizing auction design.

This same issue arises in another recent degree of complementarity definition—the supermodular (SM) degree introduced by Feige and Izsak (2013). Two items j, j' have a *supermodular* relation if there exists a (possibly empty) set $S \subseteq [m] \setminus \{j, j'\}$ such that $v(S \cup \{j, j'\}) - v(S \cup \{j'\}) > v(S \cup \{j\}) - v(S)$ (i.e., S exposes the synergy between j and j'). A valuation has supermodular degree d if each item j has at most d other items with a supermodular relation.[1] In Examples 1 and 2, the supermodular degree is at least $\max_j |S_j| - 1$ but could be larger depending on the structure in which the sets $\{S_j\}_j$ intersect.

To further emphasize the relevance of a novel measure, consider the following two instantiations of Example 2: (1) Every pair of experts can complete a project unique to them, and there is a single superteam of 10 experts that can complete its own project; (2) every set of 10 experts can complete a project unique to it. In both settings, the PH degree is 10, and the supermodular degree is m. Yet, setting (2) seems vastly more complex than (1), and an ideal measure of complementarity should capture this.

1.3. Our Notion

As mentioned, our notion begins with the PH model. The difference is in how we measure the degree of complementarity. In the language of our examples, the set function $w(\cdot)$ has $w(S_j) = v_j$ for all projects/diplomas j, and $w(T) = 0$ for all T that do not correspond to a project/diploma. We define the degree of complementarity to be the maximum over all items of the number of projects/diplomas requiring that item. Note that one simple comparison between the PH degree and our measure is that the PH degree of a hypergraph is the maximum size of any hyperedge, whereas our measure is the maximum degree of any node.

Revisiting the two objections previously raised, observe that when only the grand bundle of all items provides nonzero value, our measure assigns this valuation a complementarity degree of one, matching the intrinsic simplicity. When every pair of experts completes a unique project (and there is one superteam of size 10), our measure assigns a complementarity degree of m. When every set of 10 experts can complete a unique project, our measure assigns a complementarity degree of $\binom{m-1}{9}$. So, our measure correctly identifies that the latter setting is significantly more complex. Notice further that our measure of complementarity ranges from 1 to 2^{m-1}, whereas previous measures range from 1 to m; so, our measure has the potential to provide a much more fine-grained evaluation.

There is one more aspect to our model of complements. Even for additive buyers, multidimensional auctions are known to be intractable when values for items are correlated (Hart and Nisan 2013, Briest et al. 2015). Specifically, there exists a distribution D over $\mathbb{R}^2$ such that when a single additive buyer's valuation is drawn from D, the optimal revenue for the seller is infinite, but the revenue of the best deterministic mechanism is one. Therefore, although our model of complementarities so far is intrinsically motivated, it is intractable without further effort. A representative approach to circumvent these impossibilities for additive buyers is the *independent items* assumption: Item values are drawn independently.

With complementarities between items, there is little sense in which we can have independence across *items*, but a notion of independence still naturally folds into the model in the following way: Simply assume that the random variables $\{v_j\}_{j\in[k]}$ are independent (that is, the students' values for various *diplomas* are independent random variables). This is the natural extension of "independent items" assumed in all prior works on this topic.

To summarize in the language of our examples, we model complements among items (courses/experts) via a set-valued function $w(\cdot)$, where $w(S)$ is equal to the value derived from a diploma requiring exactly the set S of courses (or a project requiring exactly the set S of experts), and a buyer's valuation for a set T of items is $\sum_{T\subseteq S} w(T)$. Over the entire population of buyers, the values $\{w(T)\}_{T\subseteq[m]}$ are independent random variables. We now proceed to describe how we model substitutes.

1.3.1. Substitutes. Consider a student planning his or her entire course selection. Naturally, constraints such as overlapping courses, geographic constraints, prerequisites, general demands on time, etc., deem some sets of courses infeasible. One can model this with a set system $\mathcal{C} \subseteq 2^{[m]}$. A set of *courses* S is in $\mathcal{C}$ if it is feasible to construct a schedule that contains all courses in S (and $S \notin \mathcal{C}$ denotes that the courses S violate some constraint and cannot be a feasible schedule). Therefore, one could fold substitutes into the student's valuation by updating $v(S)$ to be $\max_{S' \in \mathcal{C}, S' \subseteq S}\{\sum_{S_j \subseteq S'} v_j\}$ (the maximum sum of values for all diplomas that can be achieved using courses from a feasible $S' \subseteq S$).

Our model merges the two prevailing models of complements and substitutes: Complements are captured via "diplomas" that require every course in a set, and substitutes are captured by downward-closed "feasibility constraints," which preclude the student from obtaining value from too many courses at once. So, to fully recap, there is a function $w : 2^{[m]} \to \mathbb{R}_+$ that takes as input a set S and outputs the value derived from a course/project that requires exactly the set S of items. There are also feasibility constraints $\mathcal{C}$, where $S \in \mathcal{C}$ denotes that it is feasible to simultaneously utilize all items in S. The buyer's value $v(S)$ for a set S of items is then $\max_{S' \subseteq S, S \in \mathcal{C}}\{\sum_{T \subseteq S} w(T)\}$. Among the entire population of buyers, the values $\{w(T)\}_{T \subseteq [m]}$ are drawn independently. Finally, we say that an instance has degree of complementarity d if the maximum over all items of the number of diplomas/courses requiring that item is d (formally, $\max_j |\{T, T \ni j \wedge w(T) > 0\}|$).

We now expound a more concrete example demonstrating the complementarity aspect (which does not have strict substitutes). This example also gives some intuition for the "hard" examples, where simple mechanisms cannot provide good approximation guarantees. Suppose that a random student in the population values diploma j at 2^{j+1} with probability $2^{-(j+1)}$ and zero otherwise (independently for all diplomas). The expected total value to students is then k, the total number of *diplomas*, and therefore, k is a trivial upper bound to the optimal revenue. Depending on the structure of the sets of courses required for each diploma, it may further indeed be possible to extract revenue $\Omega(k)$, even when $k = 2^{\Omega(m)}$ (if, for instance, for every set of size $m/2$ courses, there is a diploma that requires exactly that set of courses; see Proposition 5 for a complete analysis).

In this example, pricing *courses* (of which there are m) achieves revenue at best $O(m) \ll k$, and selling the grand bundle for a single price achieves revenue only $O(1)$ (see Proposition 5 for analysis of both these claims). Therefore, the better of selling items separately and selling the grand bundle, in this example, cannot achieve an approximation better than $O(\frac{m}{k})$ to the optimal revenue.

1.4. Main Result and Techniques

Our main result (Theorem 2) is that the mechanism proposed by Babaioff et al. (2014)—the better of selling separately (post a price on each item, let the buyer purchase whatever subset she likes) and bundling together (post a single price on the grand bundle, let the buyer purchase or not)—achieves a tight $\Theta(d)$ approximation whenever buyer valuations exhibit complementarity of degree at most d (by our complementarity measure—the maximum over all items of the number of diplomas/projects that require that item).

We show that our notion of complementarity best fits our model. If instead we measure complementarity via the "supermodular degree," then there exist populations in our model with supermodular degree d for which the better of selling separately and bundling together achieves only a $\Omega(2^d/d)$ approximation. Similarly, if we instead measure complementarity via the "positive hypergraph degree," then there exist populations in our model with positive hypergraph degree d for which the better of selling separately and bundling together achieves only a $\Omega(\sum_{\ell \leq d}\binom{m}{\ell}/m)$ approximation. Both notions of degree are defined formally in Section 6, where the lower bounds are proved. Our point is not that $\Theta(d)$ is a "better" bound than $\Omega(2^d/d)$ (this is arguably not a fair comparison as the measures operate on different scales) but rather, that "supermodular degree" and "positive hypergraph degree" are incapable of capturing the smooth transition from low to high degrees of complementarities as they can only take on m different values but provide guarantees that range from one to $\Omega(2^m)$. In comparison, our notion of degree of complementarity takes on 2^{m-1} different values and provides guarantees that range from one to $\Omega(2^m)$, allowing for an exponentially finer-grained trade-off.

1.4.1. Our Techniques.

Our starting point is a duality-based upper bound on the optimal achievable revenue coming from recent work of Cai et al. (2016). Their upper bound decomposes into three parts, which they call SINGLE, CORE, and TAIL. So, the goal is to show that selling separately well approximates SINGLE and that bundling together well approximates CORE and TAIL. Fortunately, the analysis of Cai et al. (2016) is fairly robust, and we are able to prove that bundling together achieves a constant factor of both CORE and TAIL via a similar approach. Our main technical contribution appears in Section 4, where we prove that selling separately gets an $O(d)$ approximation to SINGLE. Incidentally, bounding SINGLE happens to be the easiest part of the analysis in Cai et al. (2016) for additive valuations.

Without getting into details about what exactly this SINGLE term is, we can still highlight the key challenge. In the context of our hard example (where the buyer's value for diploma j is 2^{j+1} with probability 2^{-j-1}), we would like to post a different price on each diploma/project. In this example, it is even the case that the optimal "diploma/project-pricing scheme" obtains a constant-factor approximation to SINGLE. The catch is that we sell courses (items), not diploma/projects. We may wish to set drastically different prices on many different diploma/projects requiring the same course, and it is unclear that we can achieve the desired diploma/project prices by cleverly setting prices on the courses separately (in fact, it could be impossible). So, our main technical contribution is an algorithm to find a subset of diploma/projects S for which it is possible to achieve any desired diploma/project pricing on S by only posting prices on courses, and the optimal revenue from diploma/projects in S is a d approximation to the optimal *diploma/project*-pricing scheme. It turns out that the right sets of diploma/projects to search for are ones where each diploma/project requires a course *not* required by any of the other diploma/projects. We show that the number of collections with this property that is needed to partition all diploma/projects tightly characterizes the approximation guarantee of selling separately and that d such collections suffice whenever each course is required by at most d diploma/projects.

Another interesting property of our analysis worth emphasizing is the following: If our scheme chooses to sell separately, it does so by first selecting a subset of at most m diplomas/projects to target and then selecting for each diploma/project a single required item to price (and all others are offered for free). Although perhaps initially counterintuitive, such schemes are not uncommon in practice in the presence of complements. For example, many iPhone apps (which could in principle be priced) are offered for free upon purchase of an iPhone. In the context of our examples, such a scheme may involve the university offering introductory classes for free and charging only for the final course completing the diploma. In the case of a recruiter, they may charge a price to recruit a true specialist (e.g., the "computer vision expert") but offer to recruit less specialized experts for free.

1.5. Further Related Work

1.5.1. Multidimensional Auction Design. A rapidly growing body of recent literature has shown that simple mechanisms are approximately optimal in quite general settings (Chawla et al. 2007, 2010, 2015; Li and Yao 2013; Babaioff et al. 2014; Kleinberg and Weinberg 2014; Bateni et al. 2015; Rubinstein and Weinberg 2015; Yao 2015; Chawla and Miller 2016; Cai and Zhao 2017; Hart and Nisan 2017). Of these, the result most related to ours is Rubinstein and Weinberg (2015), which proves that the better of selling separately and bundling together achieves a constant-factor approximation for a single buyer whose valuation is drawn from a population that is "subadditive with independent items" (note that their approximation guarantees in this model are improved by Chawla and Miller 2016 and Cai and Zhao 2017).Their model is similar to our model with $d = 1$ (but neither subsumes the other), so our results can best be interpreted as an extension of theirs to buyers whose valuations also exhibit complementarity.

In terms of techniques, our work makes use of a recent duality framework developed in Cai et al. (2016). The same duality framework has been used in concurrent work by the present authors to prove multidimensional "Bulow–Klemperer" results (Eden et al. 2017) and independent work by others to design simple, approximately optimal auctions for multiple subadditive bidders (Cai and Zhao 2017). Still, the duality theory is only used to provide an upper bound on the revenue in all these cases, and the remaining technical contributions are disjoint. In particular, for the present paper, Section 3 has a high technical overlap with these works, and Section 5 bears some similarity. However, our main technical contribution lies in Section 4, which is unique to the problem at hand.

1.5.2. Agents with Complements. In recent years, there has also been a rapid growth in the design of algorithms and mechanisms in the presence of complements (Abraham et al. 2012; Feige and Izsak 2013; Feldman and Izsak 2014, 2017; Feige et al. 2015; Feldman et al. 2015, 2016; Nguyen et al. 2016). These works consider many different aspects: for example, assuming strategic behavior of agents (or not), assuming the existence of strict substitutes (or not), or focusing on simple mechanisms and quantifying the efficiency of equilibria. In all these works, some notion of degree of complementarity is cast on a class of valuation functions, and the approximation ratio guaranteed grows as a function of complementarity degree. It is noteworthy that quite often, different settings motivate different degrees of complementarity to best capture the degradation in possible guarantees. For instance, Abraham et al. (2012) uses the PH degree, Feige and Izsak (2013) uses the supermodular degree, Feige et al. (2015) and Feldman et al. (2015) use the maximum over positive hypergraph (MPH) degree, and Feldman et al. (2016) uses the positive supermodular degree.

Although both Abraham et al. (2012) and this paper use a positive hypergraph to describe an agent's valuations, our degree of complementarity is different than theirs. Although they set the degree of complementarity as the maximum size of a hyperedge in the hypergraph (the number of items in the largest hyperedge), our degree of complementarity is the maximum degree of an item (the number of hyperedges that contain an item). Moreover, we allow substitutability by introducing a downward-closed feasibility constraint over items, whereas Abraham et al. (2012) does not consider valuations that exhibit both complementarity and substitutability. See Section 6 for details about the different notions of complementarity used in previous works.

In comparison with the literature, ours is the first to consider *revenue* maximization for buyers with complements. Earlier work does indeed consider revenue maximization for buyers with complements (Levin 1997, Milgrom 2007, Day and Milgrom 2008) but from a fairly different perspective. For instance, Milgrom (2007) and Day and Milgrom (2008) consider core-selecting auctions in a non-Bayesian setting. Levin (1997) considers single-parameter valuations in a Bayesian setting and explicitly notes the challenges in extending to multiparameter settings. Additionally, there has also already been follow-up work following an initial announcement (one-page abstract) of portions of this work in the Conference on Economics and Computation 2019: Cai et al. (2018) study revenue maximization in a Bayesian multiparameter setting for a model of "proportional" complements. Their model is more general than ours in some ways and more restrictive in others. We expect further results along this line in the future.

1.6. Discussion and Future Work

We present the first simple and approximately optimal mechanism for a buyer whose valuation exhibits both substitutes and complements. We show that for a natural notion of "degree of complementarity," the better of selling separately and selling together achieves a tight $\Theta(d)$ approximation to the optimal revenue. We provide rigorous evidence that this notion best fits our model via large lower bounds for classes of valuations that previous definitions would "award" a low degree of complementarity.

Our main technical contribution is an algorithm to partition a collection of sets into subcollections such that each set (in the subcollection) contains an item not contained in the others (in that same subcollection). Because of the robustness of previously developed tools like the "core-tail" decomposition (Li and Yao 2013, Babaioff et al. 2014, Rubinstein and Weinberg 2015, Yao 2015, Chawla and Miller 2016, Cai and Zhao 2017) and duality-based benchmarks (Cai et al. 2016), we are able to focus our technical contributions to the specific problem at hand.

One immediate direction for future work would be to see whether simple mechanisms remain approximately optimal for *multiple* buyers with complementarity degree d. Doing so would likely require at least one substantial innovation beyond the ideas in this paper, as even the $d = 1$ case remains open (even considering the recent breakthrough result of Cai and Zhao 2017).

2. Preliminaries

We consider a setting in which a seller wishes to sell a set M of m items to a single buyer. The buyer has a valuation function v that assigns a nonnegative real number $v(S)$ to every bundle of items $S \subseteq M$. The valuation is normalized ($v(\emptyset) = 0$) and monotone ($v(S) \le v(T)$ whenever $S \subseteq T$). We slightly abuse notation and let $v(X) = \mathbb{E}_{S\sim X}[v(S)]$ when X is a random set.

2.1. Valuations with Substitutes and Complements

2.1.1. Complements. An increasingly popular model to represent complementarities is via a *positive hypergraph representation*: $v(S) = \sum_{T\subseteq S} w(T)$, where $w : 2^M \to \mathbb{R}^+$ is a nonnegative weight function. Intuitively, $w(T)$ denotes the bonus value that the buyer enjoys from owning exactly the set of items T (in addition to the value the buyer already enjoys for proper subsets of T). In the language of Section 1, $w(T)$ denotes the buyer's value for the diploma that requires exactly the courses in T. We sometimes refer to T as a *hyperedge*, thinking of $w(\cdot)$ as a weight function on the hypergraph with nodes M. We say that v (or w) exhibits complementarities *of degree* d if for every item $i \in M$, $|\{S \subseteq M : i \in S \text{ and } w(S) > 0\}| \le d$.

A simple example of a positive hypergraph representation is the following. Let v be an additive valuation; then, defining $w(\{i\}) = v(\{i\})$ and $w(T) = 0$ for every $|T| > 1$ yields $v(S) = \sum_{T\subseteq S} w(T)$.

2.1.2. Substitutes. An equally popular model to represent substitutes is via a set system capturing combinatorial constraints. Let $\mathcal{C} \subseteq 2^M$ denote a downward-closed set system over the items M; then, v assigns values to the sets in $\mathcal{C}$, and for every other set S, $v(S) = \max_{T\subseteq S, T\in\mathcal{C}}\{v(T)\}$. Intuitively, if $S \notin \mathcal{C}$, then at least some items in S are substitutes, and the buyer does not derive value from all of S.

Many valuations that exhibit only substitutabilities are representable as "additive subject to constraints" (i.e., $v(S) = \max_{T\subseteq S, T\in\mathcal{C}}\{\sum_{i\in T} v(\{i\})\}$). For example, unit-demand valuations can be represented with $\mathcal{C} = \{T \subseteq M : |T| \le 1\}$. Constraints that require a student to take at most five courses could be represented with $\mathcal{C} = \{T \subseteq M : |T| \le 5\}$. Constraints that require a student to take no more than 60 hours of coursework could be represented with $\mathcal{C} = \{T \subseteq M : \sum_{j \in T} h_j \leq 60\}$, where h_j is the number of hours of coursework for class j. Constraints that require a student to not take overlapping classes could first build a graph G with an edge between class i and j if they overlap and then define $\mathcal{C}$ to be all independent sets of G. Enforcing multiple of the aforementioned constraints simultaneously is simply a matter of taking the intersection of the defined $\mathcal{C}$s.

2.1.3. Complements and Substitutes. We choose to model substitutes and complementarities together by combining the two models. That is, there is a positive hypergraph representation w that represents complementarities, combinatorial constraints $\mathcal{C}$ that represent substitutabilities, and $v(S) = \max_{T \subseteq S, T \in \mathcal{C}} \{\sum_{U \subseteq T} w(U)\}$. We assume without loss of generality that $w(T) = 0$ for all $T \notin \mathcal{C}$.

2.2. Distributions of Valuations

We model our buyer valuation $v(\cdot)$ as being drawn from the population D in the following way. There are some constraints $\mathcal{C}$ that are fixed (not randomly drawn). Each $w(T)$ is then drawn independently from some distribution D'_T for every $T \in \mathcal{C}$, and $v(S) = \max_{T \subseteq S, T \in \mathcal{C}} \{\sum_{U \subseteq T} w(U)\}$.

We say that D has complementarity d if all valuations in the support of D have complementarity d. Note that this implies that D has complementarity d if and only if for every item $i \in M$,

$$|\{T \ni i : \Pr[w(T) = 0] < 1\}| \leq d.$$

We use V to denote the support of D, $f(v)$ to denote $\Pr_{\hat{v} \leftarrow D}[\hat{v} = v]$, and $f_T(y) = \Pr_{x \leftarrow D'_T}[y = x]$.

2.2.1. Discrete Vs. Continuous Distributions. Like Cai et al. (2016), we only explicitly consider distributions with finite support. Similar to their results, all of our results immediately extend to continuous distributions as well via a standard discretization argument (Bei and Huang 2011, Hartline et al. 2011, Daskalakis and Weinberg 2012, Hartline and Lucier 2015, Rubinstein and Weinberg 2015). We refer the reader to Cai et al. (2016) for the formal statement and proof.

Theorem 1 assumes that for every single-dimensional random variable X and number $q \in [0, 1]$, there exists a threshold p so that $X \geq p$ with probability *exactly* q, which might a priori seem problematic for discrete distributions. Fortunately, standard "smoothing" techniques allow this assumption to be valid for discrete distributions. A formal discussion of this appears in remark 2.4 of Rubinstein and Weinberg (2015).

2.3. Mechanisms

2.3.1. Truthful Mechanisms and Revenue Maximization. Formally, a mechanism $\mathcal{M}$ has two mappings $X : V \to \Delta(2^M)$ and $p : V \to \mathbb{R}$. Mapping X takes as input a valuation v and awards a (potentially random) subset of items. p takes as input a valuation v and charges a price. Mechanism $\mathcal{M}$ is then *truthful* if for all $v, v' \in V$, $v(X(v))- p(v) \geq v(X(v')) - p(v')$ (note that for a single buyer, there is no need to distinguish between Bayesian incentive compatible and dominant strategy incentive compatible—the definitions coincide). Alternatively, one can view a mechanism as a *menu* that lists options of the form (X, p), where $X \in \Delta(2^M)$ and $p \in \mathbb{R}$. A buyer with value $v(\cdot)$ then selects the menu option $\arg\max \times \{v(X) - p\}$. It is easy to see the equivalence between the two representations: Simply setting $(X(v), p(v)) = \arg\max\{v(X) - p\}$ takes one from the menu view to a truthful mechanism. We denote by REV(D) the optimal revenue attainable by any truthful mechanism when buyer valuations are drawn from the population D.

2.3.2. Simple Mechanisms. The two simple mechanisms we study are selling separately (SREV) and bundling together (BREV). We denote by BREV(D) the optimal expected revenue attainable by selling all items together and drop the parameter D when it is clear from context. It is well known that BREV(D) = $\max p \cdot \Pr[v(M) \geq p]$ (Myerson 1981). SREV requires some care as it may be computationally intractable for a buyer to even decide, given item prices, which set of items provides her the optimal utility. In such cases, it is not clear which set of items a computationally bounded buyer will purchase. Therefore, counting on the buyer to compute an optimal buying strategy may be an undesirable solution concept (from a computational perspective). In certain cases, however, one can easily determine whether item i will be purchased: For example, if *every* set that the buyer is even willing to purchase contains item i, then certainly the buyer will purchase (at least) item i. Therefore, if we only count item i as sold whenever it is contained in every set that the buyer is willing to purchase, we certainly never overestimate the revenue achieved by any rational buyer. Put another way, our only assumption on the buyer's behavior is that whenever there exists a set yielding strictly positive utility, they choose to purchase a set that yields strictly positive utility (not necessarily their utility-maximizing set)—our revenue guarantees will hold for any buyer whose behavior satisfies this property.

This is similar to the approach used by Rubinstein and Weinberg (2015): We define SREV* to be the optimal revenue attainable by any item pricing *only counting* *an item as sold if every set the buyer is willing to purchase contains that item*. More formally, for a given item pricing $\vec{p}$ and valuation v, let $I_i(\vec{p}, v) = 1$ if $\exists S \ni i, v(S) - \sum_{j\in S} p_j > 0$ and $\forall S \not\ni i, v(S) - \sum_{j\in S} p_j \le 0$, and $I_i(\vec{p}, v) = 0$ otherwise. Then, $\text{SREV}^*(D) = \max_{\vec{p}} \mathbb{E}_{v\leftarrow D}[\sum_i I_i(\vec{p}, v)\cdot p_i]$.

2.4. The Copies Environment

In our bounds, we shall make use of a related "copies environment" also utilized in Chawla et al. (2007, 2010, 2015) and Kleinberg and Weinberg (2014). For any product distribution $D' = \times_{i=1}^{k} D'_i$, we define the corresponding copies setting as follows. There is a single item for sale and k buyers. Buyer i's value for the item is drawn from the distribution D'_i. For instance, in our model, the hypergraph representation of the valuation is drawn from $D' = \times_S D'_S$, so we would have a buyer for every subset, with buyer S's value drawn from the distribution D'_S. We emphasize that in the copies setting, there is a single item for sale and a buyer for every subset S. There are no longer any feasibility constraints on which items can be simultaneously purchased: There is a single item that can be awarded to at most one buyer (or no one).

We can then define the benchmark $\text{OPT}^{\text{copies}}(D')$ to be the expected revenue obtained by the optimal mechanism of Myerson (Myerson 1981) on input D'. Note that this is equal to $\mathbb{E}_{w\leftarrow D'}[\max_T\{\bar{\varphi}_T(w(T)), 0\}]$, where $\bar{\varphi}_T(\cdot)$ denotes Myerson's *ironed virtual value* for the distribution D'_T. We make use of the following result, the proof of which appears for completeness in Appendix B.

Theorem 1 (Chawla et al. 2010)**.** *For any $q \le 1$, there exist (possibly random) prices $\{p_T\}_T$ such that*

1. *Revenue is high: $OPT^{copies}(D') \le \frac{1}{q}\sum_{T\subseteq M} \mathbb{E}_{p_T}[p_T \cdot \Pr_{x\leftarrow D'_T}[x \ge p_T]]$.*
2. *Probability of sale is low: $\sum_{T\subseteq M} \mathbb{E}_{p_T}[\Pr_{x\leftarrow D'_T}[x \ge p_T]] \le q$.*
3. *Moreover, each p_T takes on at most two values. If D'_T is regular, then p_T is a point mass.*

3. Our Duality Benchmark and Main Theorem Statement

We extend the duality framework of Cai et al. (2016) to our setting in a natural manner. Full technical details are deferred to Appendix A. The only technical detail needed for stating our revenue benchmark is the following: We partition the valuation space V into $2^m - 1$ different regions, depending on which hyperedge is the most valuable to a buyer with valuation v. Specifically, we say that v is in region R_A if $A = \arg\max_{T\subseteq M}\{w(T)\}$, with ties broken lexicographically.

Corollary 1. *For valuation distribution D established by drawing a hypergraph representation $w \leftarrow \prod_S D'_S$ and returning $v(S) = \max_{T\subseteq S, T\in\mathcal{C}}\{\sum_{U\subseteq T} w(U)\}$,*

$$REV(D) \le \mathop{\mathbb{E}}_{v\leftarrow D}\left[\max_{S\in\mathcal{C}}\left\{\sum_{T\subseteq S} w(T)\cdot \mathbb{1}[v \notin R_T]\right\}\right] \qquad (NONFAVORITE)$$

$$+ \mathop{\mathbb{E}}_{v\leftarrow D}\left[\sum_{S\subseteq M} \max\{0, \bar{\varphi}_S(w(S))\}\cdot \mathbb{1}[v \in R_S]\right]. \qquad (SINGLE)$$

We defer to Appendix A any discussion of how this benchmark is derived but provide here some intuition to help parse it. Corollary 1 upper bounds the revenue with two terms. The second term, SINGLE, sums over all S a term that is zero whenever $v \notin R_S$ (that is, S is not the buyer's favorite diploma). When $v \in R_S$ (that is, S is the buyer's favorite diploma), it sums the Myersonian (ironed) virtual value for diploma S (as defined by distribution D'_S). The first term, NONFAVORITE, is simply the buyer's value for the grand bundle of all items *only counting contributions from diplomas that are not their favorite*.

In Section 4, we show that $\max\{\text{SREV}^*, \text{BREV}\}$ gets a $4(d+1)$ approximation to SINGLE (Proposition 1). This portion of the analysis develops techniques specific to buyers with restricted complements. In Section 5, we show that BREV gets a 12 approximation to NONFAVORITE (Proposition 4). This portion of the analysis will look somewhat standard to the reader familiar with Cai et al. (2016), with a little extra work to extend their main ideas to our setting. We conclude this section with our main theorem

Theorem 2. *For a distribution D that has complementarity d, $REV \le (4d + 16)\max\{BREV, SREV^*\}$.*

Proof. Combine Propositions 1 and 4 with Corollary 1 to get

$$\begin{aligned}\text{REV}(D) &\le 4d\text{SREV}^*(D) + 16\text{BREV}(D)\\ &\le (4d+16)\max\{\text{SREV}^*, \text{BREV}\}. \quad \square\end{aligned}$$

4. Bounding SINGLE

In this section, we show that the better of selling items separately and selling the grand bundle gets an $O(d)$ approximation to SINGLE. Specifically, we prove the following.

Proposition 1. *$SINGLE \le 4dSREV^* + 4BREV$.*

4.1. Relating SINGLE to $\text{OPT}^{\text{copies}}$

We begin by relating SINGLE to $\text{OPT}^{\text{copies}}$.

Observation 1. SINGLE $\leq$ OPT$^{\text{copies}}$.

Proof. First, observe that there is exactly one S for which $\mathbb{1}[v \in R_S] = 1$. So, it is certainly the case that for all v (with $v(S) = \sum_{T \subseteq S} w(T)$), we have

$$\sum_{S \subseteq M} \max\{0, \bar{\varphi}_S(w(S))\} \cdot \mathbb{1}[v \in R_S] \leq \max_{S \subseteq M}\{0, \bar{\varphi}_S(w(S))\}.$$

$$\Rightarrow \mathop{\mathbb{E}}_{v \leftarrow D}\left[\sum_{S \subseteq M} \max\{0, \bar{\varphi}_S(w(S))\} \cdot \mathbb{1}[v \in R_S]\right] \leq \mathop{\mathbb{E}}_{v \leftarrow D}\left[\max_{S \subseteq M}\{0, \bar{\varphi}_S(w(S))\}\right].$$

The left-hand side is exactly SINGLE, and the right-hand side is exactly OPT$^{\text{copies}}$. □

Note that if the buyer's valuation was additive, at this point we would already be finished. We could simply set the prices guaranteed by Theorem 1 and be done. As we consider more complex buyer valuations, there are two barriers we must overcome. The first is because of substitutability: If we try to set prices on each subset separately, just because the buyer is *willing* to purchase set S does not mean he will *choose* to purchase set S because he may purchase some substitutes instead. Note that this issue does not arise in absence of substitutes: If the buyer is willing to purchase S by itself, he is certainly willing to add S to any other set of purchased items. The second barrier is because of complementarity: Even after we decide the "correct" price to charge for set S, we can only set prices on *items* and not on *bundles*. Therefore, the prices we want to set for different bundles necessarily interfere with each other. This is the novel barrier unique to values with complementarity and is also the only part of the analysis where the (necessary) factor of d arises.

4.2. Overcoming the Complements and Substitutes Barriers

The first step to overcoming the complements barrier is to find a subset of bundles for which we can still set the appropriate prices. As a warm-up, let us see what the argument would look like assuming that there were only complements and no substitutes ($\mathcal{C} = 2^M$).

Lemma 1. *Let $\mathcal{C} = 2^M$ and $T_1, \ldots, T_k$ be subsets of M such that $T_i \not\subseteq \cup_{j \neq i} T_j$ for all i. Then, for all $\{p_T\}_{T \subseteq M}$, SREV $\geq \sum_i p_{T_i} \Pr_{x \leftarrow D'_{T_i}}[x \geq p_{T_i}]$.*

Proof. Set price p_{T_i} on the item contained in T_i but not $\cup_{j \neq i} T_j$ (if there are multiple, select one arbitrarily). Then, by hypothesis, the price the bidder would have to pay in order to receive the entire set T_i is exactly p_{T_i}. Because $\mathcal{C} = 2^M$, whenever $w(T_i) \geq p_{T_i}$, the buyer will choose to purchase the set T_i in addition to whatever else he or she chooses to purchase. Therefore, the item contained in T_i but not $\cup_{j \neq i} T_j$ is purchased with probability at least $\Pr_{x \leftarrow D_{T_i}}[x \geq p_{T_i}]$, and the revenue of this item pricing is at least $\sum_i p_{T_i} \Pr_{x \leftarrow D'_{T_i}}[x \geq p_{T_i}]$. □

The proof of Lemma 1 makes use of the assumption that $\mathcal{C} = 2^M$ in exactly one place: to argue that whenever $w(T_i) \geq p_{T_i}$, the buyer chooses to purchase the complete set T_i. When $\mathcal{C} \neq 2^M$, it may be the case that even though the buyer is willing to purchase set T_i, she chooses to purchase substitutes instead. We can remove this assumption on $\mathcal{C}$ by restricting attention to certain price vectors.

Lemma 2. *Let $\mathcal{C}$ be any downward-closed set system and $T_1, \ldots, T_k$ be subsets of M such that $T_i \not\subseteq \cup_{j \neq i} T_j$ for all i. Then, for all $\{p_T\}_{T \subseteq M}$ such that $p_T \geq 4$BREV for all T, SREV$^* \geq \frac{1}{4} \sum_i p_{T_i} \Pr_{x \leftarrow D'_{T_i}}[x \geq p_{T_i}]$.*

Proof. Set price $p_{T_i}/2$ on the item contained in T_i but not $\cup_{j \neq i} T_j$ (if there are multiple, again select one arbitrarily). The price the bidder would have to pay in order to receive the entire set T_i is exactly $p_{T_i}/2$. Suppose $w(T_i) \geq p_{T_i}$. Then, the buyer is not only willing to purchase T_i but also gets utility at least $p_{T_i}/2$ for doing so. The only reason she would choose not to purchase this set is if there was some other set S with $T_i \not\subseteq S$ and $v(S) \geq p_{T_i}/2 \geq 2$BREV. As $v(S) \leq v(M) - w(T_i)$ for all such S, in order for such a set to exist, it must be the case that $v(M) - w(T_i) \geq 2$BREV. Clearly, this occurs with probability at most $\frac{1}{2}$, as otherwise we could set price 2BREV on the grand bundle, sell with probability strictly larger than $\frac{1}{2}$, and make revenue strictly larger than BREV. Moreover, $v(M) - w(T_i) = \sum_{U \neq T_i} w(U)$ is completely independent of $w(T_i)$. Therefore, even conditioned on $w(T_i) \geq p_{T_i}$, the probability that the bidder is interested in some other set S with $T_i \not\subseteq S$ is at most $\frac{1}{2}$, and therefore, the buyer indeed chooses to purchase T_i with probability at least $\Pr_{x \leftarrow D'_{T_i}}[x \geq p_{T_i}] \cdot \frac{1}{2}$. □

Finally, we can combine Lemma 2 with Theorem 1 to reduce our search to the problem of partitioning the hyperedges into collections $H_x = \{T_{x1}, \ldots, T_{xk_x}\}$ such that $T_{xi} \not\subseteq \cup_{j \neq i} T_{xj}$ for all i.

Corollary 2. *Let $\mathcal{C}$ be any downward-closed set system, and let $\{H_x\}_{x \in [k]}$ be a partition of the hyperedges $\{T : f_T(0) < 1\}$ such that for all x, and all $T \in H_x$, $T \not\subseteq \cup_{T' \in H_x \setminus \{T\}} T'$. Then, $4k$SREV$^* + 4$BREV $\geq$ SINGLE.*

Proof. Take $q = 1$ in Theorem 1, and let $\{p_T\}_{T \subseteq M}$ be the guaranteed (randomized) prices. By Theorem 1, condition 3, there exist two deterministic prices $p_T^H \geq p_T^L$ and probabilities q_T such that $p_T = p_T^H$ with probability q_T

and $p_T = p_T^L$ with probability $1 - q_T$. Therefore, Theorem 1, condition 1 can be rewritten as

$$\mathrm{OPT}^{\text{copies}} \le \sum_{T\subseteq M} q_T p_T^H \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^H\right] + (1-q_T)p_T^L \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^L\right].$$

We can further rewrite this by breaking up the two sums into prices that exceed 4BREV and those that do not; let $\mathcal{B} = 4\mathrm{BREV}$ for simplicity:

$$\begin{aligned}\mathrm{OPT}^{\text{copies}} \le &\sum_{T\subseteq M, p_T^H \le \mathcal{B}} q_T p_T^H \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^H\right] \\ &+ \sum_{T\subseteq M, p_T^L \le \mathcal{B}} (1-q_T)p_T^L \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^L\right] \\ &+ \sum_{T\subseteq M, p_T^H > \mathcal{B}} q_T p_T^H \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^H\right] \\ &+ \sum_{T\subseteq M, p_T^L > \mathcal{B}} (1-q_T)p_T^L \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^L\right].\end{aligned}$$

By condition 2 of Theorem 1, we have

$$\sum_{T\subseteq M} q_T \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^H\right] + (1-q_T) \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T^L\right] \le 1.$$

Therefore, as all prices in the top sum are at most $\mathcal{B}$, the entire top two terms sum to at most $\mathcal{B} = 4\mathrm{BREV}$.

For the bottom two terms, there is no term for T if $p_T^H \le \mathcal{B}$. If $p_T^H > \mathcal{B} \ge p_T^L$, define $p_T = p_T^H$. If $p_T^H > p_T^L > \mathcal{B}$, then set p_T to whichever of $\{p_T^H, p_T^L\}$ maximizes $p_T \cdot \Pr_{x\leftarrow D'_T}[x \ge p_T]$. Then, $\sum_{T\subseteq M, p_T^H>\mathcal{B}} p_T \cdot \Pr_{x\leftarrow D'_T}[x \ge p_T]$ is at least as large as the bottom two terms. Moreover, as all $p_T > \mathcal{B}$, we can apply Lemma 2 to conclude that for all $T_1, \dots, T_k$ uch that $T_i \not\subseteq \cup_{j\ne i} T_j$ or all i, $\mathrm{SREV}^* \ge 1/4 \sum_i p_{T_i} \Pr_{x\leftarrow D'_{T_i}}[x \ge p_{T_i}]$.

Finally, as $\{H_x\}_{x\in[k]}$ partitions the hyperedges so that for all x and $T \in H_x$, $T \not\subseteq \cup_{T'\in H_x\setminus\{T\}} T'$, we get

$$\sum_{T\subseteq M, p_T^H > \mathcal{B}} p_T \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T\right] = \sum_{x=1}^{k} \sum_{T\in H_x, p_T^H>\mathcal{B}} p_T \cdot \Pr_{x\leftarrow D'_T}\left[x \ge p_T\right] \le 4k \cdot \mathrm{SREV}^*.$$

The last inequality is because of Lemma 2 and completes the proof.

So, the last remaining task is to find a good partition of hyperedges, such that within each partition, every hyperedge contains at least one item not contained in the other hyperedges in the same partition. We isolate this contribution in Section 4.3.

4.3. Partitioning Hyperedges with Restricted Complements

We provide a high-level description of our algorithm here and give pseudocode in Figure 1. Recall that the algorithm takes as input a set of hyperedges and returns a partition of the hyperedges $\{H_x\}_x$, so that in each partition H_x, every hyperedge $S \in H_x$ contains an item that is not in any other hyperedge $T \in H_x$. The algorithm iteratively constructs each H_x and initially initializes H_x to contain all remaining hyperedges. Then, it iteratively eliminates all "bad" hyperedges (those that *do not* contain an item absent from the others) until the remaining hyperedges have the desired property. In the proof of Theorem 3, it is easy to show that the algorithm outputs a feasible partition, and the trick is guaranteeing that each iteration makes sufficient progress toward finalizing the partition.

Theorem 3. *For any set of hyperedges $E \subseteq 2^M$, the algorithm in Figure 1 returns a partition of $E = \{H_x\}_{x\in[k]}$ such that*

1. *For all x, and all $T \in H_x$, $T \not\subseteq \cup_{T'\in H_x\setminus\{T\}} T'$.*
2. *$k \le \max_i\{|\{T \in E : i \in T\}|\}$.*

Proof. First, it is clear that the algorithm indeed properly outputs a partition of E: Observe that because of line 2d in Figure 1, when a hyperedge is permanently assigned to some E_i, it will not be assigned to any $E_{i'}$, which implies that all the E_is are disjoint. Also, every hyperedge is either permanently assigned to some E_i or remains in E_{curr}, which by line 2, implies that the algorithm terminates only when every hyperedge is permanently assigned to some E_i. So, every hyperedge is contained in some partition, and the partitions are disjoint.

That the output partition satisfies Property 1 is easy to verify: For any x, $T \in H_x$ only the check in 2c passes

Figure 1. An Edge-Partitioning Process

Partition-Edges
Input: List of hyperedges, $E \subseteq 2^M$.
Output: A partition of E into $\{H_x\}_x$ such that for all x and all $T \in H_x$, $T \not\subseteq \cup_{T'\in H_x\setminus\{T\}} T'$.

1. $E_{\text{curr}} \leftarrow E$, $i \leftarrow 0$.
2. While $E_{\text{curr}} \ne \emptyset$:
   (a) $i \leftarrow i + 1$
   (b) $E_i \leftarrow E_{\text{curr}}$.
   (c) For each $T \in E_i$ (in arbitrary order): If $T \subseteq \bigcup_{S\in E_i\setminus\{T\}} S$ Then $E_i \leftarrow E_i \setminus \{T\}$.
   (d) $E_{\text{curr}} \leftarrow E_{\text{curr}} \setminus E_i$.
3. Return the partition $\{E_j\}_{j\in[i]}$.

for T and (the present) H_x. After the check passes, some other edges will be removed from H_x before the output. Clearly, removing edge from H_x cannot cause T to all of a sudden be contained in $\cup_{T'\in H_x\setminus\{T\}}T'$ when it was previously not contained. So, Property 1 is satisfied.

To prove Property 2, first denote by E^i_{curr} the state of E_{curr} at the start of iteration i. We will show that $\cup_{T\in E_i}T = \cup_{T\in E^i_{\text{curr}}}T$. In other words, every element contained in some hyperedge in E^i_{curr} is still contained in some hyperedge in E_i. To see this, observe that when E_i is first set to E^i_{curr}, we clearly have $\cup_{T\in E_i}T = \cup_{T\in E^i_{\text{curr}}}T$. The only time hyperedges are removed from E_i is in step 2c. Note that in order for a hyperedge to be removed from E_i, it must be the case that $T \subseteq \cup_{T'\in E_i\setminus\{T\}}T'$. In other words, in order to remove T from E_i, it must be that all the elements contained in T are also contained in $\cup_{T'\in E_i\setminus\{T\}}T'$. Therefore, removing T does not change $\cup_{T'\in E_i}T'$, and when we terminate, we maintain $\cup_{T\in E_i}T = \cup_{T\in E^i_{\text{curr}}}T$.

To see why this implies Property 2, note that it implies that if for any i, $|\{T \in E, i \in T\}| = d$, then i will be contained in at least one hyperedge in all of $E_1, \dots, E_d$, and therefore, no hyperedges containing i remain in E^{d+1}_{curr}. In particular, for $d = \max_i\{|\{T \in E, i \in T\}|\}$, it is the case that for all i, no hyperedges containing i remain in E^{d+1}_{curr}, and therefore, the algorithm terminates with at most d partitions. □

We can now combine everything to provide a proof of Proposition 1.

Proof of Proposition 1. Combining Theorem 3 with Corollary 2, we get that whenever D has complementarity d, that $4d\text{SREV}^* + 4\text{BREV} \geq$ SINGLE, completing the proof.

5. Bounding NONFAVORITE

In this section, we bound NONFAVORITE using similar ideas to those developed in Cai et al. (2016). Much of the process will look familiar to experts familiar with Rubinstein and Weinberg (2015) and Cai et al. (2016), but there are a couple of new ideas sprinkled in. We begin by breaking NONFAVORITE into CORE + TAIL, as is by now standard (t will be chosen later).

Lemma 3. *NONFAVORITE is upper bounded by the following:*

$$\mathop{\mathbb{E}}_{v\leftarrow D}\left[\max_{S\in\mathcal{C}}\left\{\sum_{T\subseteq S} w(T)\cdot \mathbb{1}[w(T)\leq t]\right\}\right] + \qquad (CORE)$$

$$\mathop{\mathbb{E}}_{v\leftarrow D}\left[\sum_{S:w(S)>t} w(S)\cdot \mathbb{1}[v\notin R_S]\right]. \qquad (TAIL)$$

Proof. The proof follows from the following algebra:

(NONFAVORITE)

$$\begin{aligned}
&= \mathop{\mathbb{E}}_{v\leftarrow D}\left[\max_{S\in\mathcal{C}}\left\{\sum_{T\subseteq S} w(T)\cdot \mathbb{1}[v\notin R_T]\right\}\right]\\
&= \mathop{\mathbb{E}}_{v\leftarrow D}\left[\max_{S\in\mathcal{C}}\left\{\sum_{T\subseteq S} w(T)\cdot \mathbb{1}[w(T)\leq t]\cdot \mathbb{1}[v\notin R_T]\right.\right.\\
&\qquad \left.\left. + w(T)\cdot \mathbb{1}[w(T)>t]\cdot \mathbb{1}\left[v\notin R_T\right]\right\}\right]\\
&\leq \mathop{\mathbb{E}}_{v\leftarrow D}\left[\max_{S\in\mathcal{C}}\left\{\sum_{T\subseteq S} w(T)\cdot \mathbb{1}[w(T)\leq t]\right\}\right]\\
&\qquad + \mathop{\mathbb{E}}_{v\leftarrow D}\left[\max_{S\in\mathcal{C}}\left\{\sum_{T\subseteq S} w(T)\cdot \mathbb{1}[w(T)>t]\cdot \mathbb{1}[v\notin R_T]\right\}\right]\\
&\leq \mathop{\mathbb{E}}_{v\leftarrow D}\left[\max_{S\in\mathcal{C}}\left\{\sum_{T\subseteq S} w(T)\cdot \mathbb{1}[w(T)\leq t]\right\}\right]\\
&\qquad + \mathop{\mathbb{E}}_{v\leftarrow D}\left[\sum_{T|w(T)>t} w(T)\cdot \mathbb{1}[v\notin R_T]\right\}\right]. \quad \square
\end{aligned}$$

5.1. Bounding CORE

Our main approach to bound CORE is to apply the same concentration bound of Schechtman (Schechtman 2003) used in Rubinstein and Weinberg (2015). Essentially, we just have to show that our valuation functions are "subadditive over independent items" for the appropriate definition of "items" (which happens to be hyperedges). It is perhaps not obvious that our valuation functions are subadditive over independent "items," but indeed, they are.

Let us first recall the definition of subadditive over independent items. In the definition, we intentionally write N instead of M to denote the set of items as the "items" in the definition may be different than the items for sale.

Definition 1. A distribution D over valuation functions $v: 2^N \to \mathbb{R}$ is *subadditive over independent items* if the following conditions hold.

1. *No externalities and independence across items.* For every item i, let Ω_i be a compact subset of a normed space (i.e., $\Omega_i = [0,1]$). There exists a product distribution D' over $\times_{i\in N}\Omega_i$ (that is, $D' = \prod_{i\in N} D'_i$) and a collection of deterministic functions $V_{\mathcal{S}}: \times_{i\in\mathcal{S}}\Omega_i \to \mathbb{R}$ such that a sample v from D can be drawn by sampling $\vec{x} \leftarrow D'$ and defining $v(S) = V_S(\vec{x}_S)$.

2. *Monotonicity.* Every v in the support of D is monotone (i.e., $v(\mathcal{S}) \leq v(\mathcal{S}')$ for every $\mathcal{S} \subseteq \mathcal{S}'$).

3. *Subadditivity.* Every v in the support of D is subadditive (i.e., $v(\mathcal{S} \cup \mathcal{S}') \leq v(\mathcal{S}') + v(\mathcal{S}')$ for all S, S').

Definition 2. Let D denote a distribution over valuation functions, D' denote the product distribution, and $\{V_S(\cdot)\}$ be the deterministic functions that witness D as subadditive over independent items. D is *c-Lipschitz* if for all $\vec{x}, \vec{y}$, and sets of items S, T, we have

$$|V_S(\vec{x}_S) - V_T(\vec{y}_T)| \leq c \cdot (|X \cup Y| - |X \cap Y| + |\{i \in X \cap Y : x_i \neq y_i\}|).$$

We use the following lemma and corollary (of a concentration inequality because of Schechtman (Schechtman 2003)) from Rubinstein and Weinberg (2015) (the bound in Corollary 3 is slightly improved from Rubinstein and Weinberg (2015), so we include a proof).

Lemma 4 (Rubinstein and Weinberg 2015). *Let D be a distribution that is subadditive over independent hyperedges, where for each hyperedge T, $v(\{T\}) \in [0, c]$ with probability one. Then, D is c-Lipschitz.*

Corollary 3 (Rubinstein and Weinberg 2015). *Suppose that D is a distribution that is subadditive over independent hyperedges and c-Lipschitz. If a is the median of $v(N)$, then $\mathbb{E}[v(N)] \leq 3a + c \cdot (2 + 1/\ln 2)$.*

Proof. By corollary 12 in Schechtman (2003), we know that for all $k > 0$,

$$\Pr[v(N) \geq 3 \cdot a + k \cdot c] \leq \min\{1, 4 \cdot 2^{-k}\}. \tag{1}$$

Substituting $x = 3 \cdot a + k \cdot c$ gets $k = (x - 3a)/c$. Therefore, Equation (1) becomes meaningful only when $4 \cdot 2^{-k} \leq 1$ (i.e., when $x \geq 2c + 3a$). Computing the expected value of $v(N)$ gives

$$\int_0^\infty \Pr[v(N) > x]dx \leq \int_0^\infty \min\{1, 4 \cdot 2^{(3a-x)/c}\}dx = 2c + 3a + 4 \cdot 2^{3a/c} \cdot \int_{2 \cdot c + 3a}^\infty 2^{-x/c} \cdot dx.$$

Computing the integral gives $-\frac{c}{\ln 2}[2^{-x/c}]_{2c+3a}^\infty = \frac{c}{\ln 2} \cdot 2^{-\frac{2c+3a}{c}} = \frac{c}{4 \ln 2} \cdot 2^{-3a/c}$, which plugged back to the equation, concludes that

$$\mathbb{E}v(N) \leq 2c + 3a + \frac{c}{\ln 2},$$

as desired. □

Finally, we just need to relate CORE to a random variable that is subadditive over independent items.

Lemma 5. *CORE is the expectation of a random variable $v_{CORE}(N)$, where $v_{CORE}(\cdot)$ is t-Lipschitz and subadditive over independent items $N = 2^M$. Moreover, $v_{CORE}(N)$ is stochastically dominated by $v(M)$.*

Proof. Let the "items" $N = 2^M$. Let the distributions $\hat{D}_T = D'_T \cdot \mathbb{1}[w(T) \leq t]$ (that is, a random variable drawn from $\hat{D}_T$ can be coupled with the random variable $w(T) \cdot \mathbb{1}[w(T) \leq t]$). Define constraints $\mathcal{C}' \subseteq 2^N$ $(= 2^{2^M})$ so that a subset U of 2^M is in $\mathcal{C}'$ if and only if there exists a set $C \in \mathcal{C}$ with $\cup_{T \in U} T \subseteq C$. In other words, $U \in \mathcal{C}'$ if and only if the union of elements of U is contained in some set in $\mathcal{C}$. Finally, define $V_U(\vec{x}_U) = \max_{U' \subseteq U, U' \in \mathcal{C}'}\{\sum_{T \in U} x_T\}$.

It is easy to see that $v_{\text{CORE}}(\cdot)$ has no externalities and independent items. It is also easy to see that $v_{\text{CORE}}(\cdot)$ is monotone. Finally, we will prove that $v_{\text{CORE}}(\cdot)$ is subadditive by observing that $\mathcal{C}'$ is downward closed. To see this, simply observe that if $U' \subseteq U$, and $\cup_{T \in U} T \subseteq C$, then clearly $\cup_{T \in U'} T \subseteq C$. So, if $C \in \mathcal{C}$ witnesses that $U \in \mathcal{C}'$ and $U' \subseteq U$, then C also witnesses that $U' \in \mathcal{C}'$.

Now that $\mathcal{C}'$ is downward closed, it is easy to see (and well known) that v_{CORE} is subadditive. For any U, W, let $X = \arg\max_{X' \subseteq U \cup W, X \in \mathcal{C}'}\{\sum_{T \in X} x_T\}$. Then, let $U' = X \cap U$ and $W' = X \cap W$. Clearly, $\sum_{T \in X} x_T \leq \sum_{T \in U'} x_T + \sum_{T \in W'} x_T$. As $\mathcal{C}'$ is downward closed, $U' \in \mathcal{C}'$ and $W' \in \mathcal{C}'$. Therefore, $v_{\text{CORE}}(W) + v_{\text{CORE}}(U) \geq \sum_{T \in U'} x_T + \sum_{T \in W'} x_T \geq \sum_{T \in X} x_T = v_{\text{CORE}}(U \cup W)$, and $v_{\text{CORE}}(\cdot)$ is subadditive.

So, finally, it remains to show that $v_{\text{CORE}}(N)$ is stochastically dominated by $v(M)$. Couple the random variable x_T drawn from $\hat{D}_T$ so that $x_T = w(T) \cdot \mathbb{1}[w(T) \leq t]$. Now consider $U^* = \arg\max_{U \subseteq 2^M, U \in \mathcal{C}'}\{\sum_{T \in U} x_T\}$. Then, we have $v_{\text{CORE}}(N) = \sum_{T \in U^*} x_T$. By definition of $\mathcal{C}'$, there exists some $C \in \mathcal{C}$ such that $T \subseteq C$ for all $T \in U^*$. Therefore,

$$\begin{aligned} v_{\text{CORE}}(N) &= \sum_{T \in U^*} x_T \leq \sum_{T \subseteq C} x_T \leq \sum_{T \subseteq C} w(T) \\ &\qquad \text{(because } x_T \leq w(T)\text{)} \\ &\leq \max_{S \subseteq M, S \in \mathcal{C}} \left\{\sum_{T \subseteq S} w(T)\right\} \\ &\qquad \text{(because } C \in \mathcal{C}\text{)} = v(M). \end{aligned}$$

So, when x_T and $w(T)$ are coupled in this way, we have $v_{\text{CORE}}(N) \leq v(M)$, and therefore, $v(M)$ stochastically dominates $v_{\text{CORE}}(N)$. □

Now, Lemma 5 combined with Corollary 3 states that $3 \cdot v(M)$ exceeds CORE $- \ t \cdot (2 + 1/\ln 2)$ with probability at least 1/2, allowing us to conclude with the following proposition.

Proposition 2. *CORE $\leq$ 6BREV $+ t \cdot (2 + 1/\ln 2)$.*

Proof. Let a be the median of the random variable $v_{\text{CORE}}(N)$. Then, $\Pr[v_{\text{CORE}}(N) \geq a] = 1/2$. As $v(M)$ stochastically dominates $v_{\text{CORE}}(N)$, we have $\Pr[v(M) \geq a] \geq 1/2$. Moreover, by Corollary 3, the fact that

CORE $= \mathbb{E}[v_{\text{CORE}}(N)]$ and that v_{CORE} is t-Lipschitz and subadditive over independent items, we have

$$\text{CORE} \le 3a + t(2 + 1/\ln 2).$$

Moreover, as $\Pr[v(M) \ge a] \ge 1/2$, we have

$$\text{BREV} \ge a/2.$$

Combining the two equations proves the proposition. □

5.2. Bounding TAIL

Our approach to bound TAIL is again similar to Cai et al. (2016). We begin by rewriting TAIL using linearity of expectation and the fact that the hypergraph representation w of valuation v is drawn from D', which is a product distribution:

$$\begin{aligned}
\text{TAIL} &= \mathop{\mathbb{E}}_{v \leftarrow D}\left[\sum_{T \subseteq M, w(T) > t} w(T) \cdot \mathbb{1}[v \notin R_T]\right] \\
&= \mathop{\mathbb{E}}_{v \leftarrow D}\left[\sum_{T \subseteq M, w(T) > t} w(T) \cdot \mathbb{1}[\exists T', w(T') > w(T)]\right] \\
&= \sum_{T \subseteq M} \mathop{\mathbb{E}}_{v \leftarrow D}[w(T) \cdot \mathbb{1}[w(T) > t \wedge v \notin R_T]] \\
&\qquad \text{(by linearity of expectation)} \\
&= \sum_{T \subseteq M} \sum_{x > t, f_T(x) > 0} x \cdot f_T(x) \cdot \Pr_{D_{-T}}[\exists T', w(T') > x] \\
&\qquad \text{(by independence across hyperedges).}
\end{aligned}$$

From here, we use essentially the same lemma from Cai et al. (2016). We have replaced their SREV with BREV, but the proof is identical.

Lemma 6 (Cai et al. 2016). *For all x, T, $x \cdot \Pr_{w \leftarrow D'_{-T}}[\exists T', w(T') > x] \le BREV$.*

Proof. For any x, we can set price x on the grand bundle. It will sell with probability at least $\Pr_{w \leftarrow D'_{-T}}[\exists T', w(T') > x]$ as whenever there is a single hyperedge with contribution x, certainly the buyer's value for the grand bundle is at least x. Therefore, BREV $\ge x \cdot \Pr_{w \leftarrow D'_{-T}}[\exists T', w(T') > x]$. □

Proposition 3. *TAIL $\le (\sum_{T \subseteq M} \Pr[w(T) > t]) \cdot BREV$.*

Proof. By Lemma 6, we get

$$\begin{aligned}
\sum_{T \subseteq M} \sum_{x > t, f_T(x) > 0} x \cdot f_T(x) \cdot \Pr_{w \leftarrow D'_{-T}}[\exists T', w(T') > x] &\le \sum_{T \subseteq M} \sum_{x > t, f_T(x) > 0} f_T(x) \cdot \text{BREV} \\
&= \sum_{T \subseteq M} \Pr[w(T) > t] \cdot \text{BREV}. \quad \square
\end{aligned}$$

5.3. Setting the Cutoff

Finally, we just need an appropriate choice of t. We will choose to set t such that $\sum_{T \subseteq M} \Pr[w(T) > t] = k$ for the appropriate choice of k. We first show how to relate t to BREV. Lemma 7 is well known, but we provide a proof for completeness.

Lemma 7. *Let $E_1, \ldots, E_k$ be independent events such that $\sum_i \Pr[E_i] = k$. Then, $\Pr[\cup_i E_i] \ge 1 - e^{-k}$.*

Proof. By independence,

$$\Pr[\cup_i E_i] = 1 - \prod_i (1 - \Pr[E_i]).$$

So, if we define $q_i = \Pr[E_i]$, we want to maximize $\prod_i (1 - q_i)$ subject to $\sum_i q_i = k$. Using a Lagrangian multiplier of λ on the constraint $\sum_i q_i = k$, we get a new objective of

$$\prod_i (1 - q_i) + \lambda \cdot \left(\sum_i q_i\right) - \lambda k.$$

We see that the partial with respect to q_i is exactly $-\prod_{j \neq i}(1 - q_j) + \lambda$. So, setting $q_i = k/n$ for all i, and $\lambda = (1 - k/n)^{n-1}$, we get that $\sum_i q_i = k$, and the partial of the Lagrangian with respect to q_i is zero for all i. Therefore, this is the optimal solution. At $q_i = k/n$ for all i, we have $\prod_i (1 - k/n) = (1 - k/n)^n \le e^{-k}$. □

Corollary 4. *If t is such that $\sum_{T \subseteq M} \Pr[w(T) > t] = k$, then BREV $\ge (1 - e^{-k})t$.*

Proof. Apply Lemma 7 to the events $E_T = \{w(T) > t\}$. Then, the probability that there exists some hyperedge T with $w(T) > t$ is at least $(1 - e^{-k})$. So, the grand bundle will sell at price t with probability at least $(1 - e^{-k})$. □

We can now complete our bound for NONFAVORITE and the proof of Theorem 2.

Proposition 4. *NONFAVORITE $\le 12BREV$.*

Proof. Combine Propositions 2 and 3, taking t such that $\sum_T \Pr[w(T) > t] = 1.66$. □

6. Lower Bounds

The following proposition shows that the factor d approximation (established in Theorem 2) is tight (up to a constant factor), even when there are no substitutes ($\mathcal{C} = 2^M$), and $d = m^{O(1)}$. The same proposition shows that our d approximation is tight up to a logarithmic factor for all d. The construction is based on a construction from Hart and Nisan (2017) used to show that BREV may be a factor of m smaller than SR EV for additive buyers, which has also inspired similar constructions (e.g., Dughmi et al. 2014).

Proposition 5. *For all $k \geq 1$, there exists a distribution D with complementarity $d \leq \binom{m}{k}$, for which*

$$REV \geq \frac{d}{2k} \max\{BREV, SREV\}.$$

Proof. Consider an integer a and a set of hyperedges E. Index the hyperedges with integers in increasing order of size with $\{1+a, 2+a, \ldots, |E|+a\}$ (we abuse notation and use e both for index and hyperedge; i.e., set of items). The product distribution D' has $f_e(0) = 1$ for all $e \notin E$, and for every $e \in E$, set $f_e(0) = 1 - 2^{-e}$ and $f_e(2^e) = 2^{-e}$. Let D be the distribution that samples $w \leftarrow D'$ and returns $v(S) = \sum_{T \subseteq S} w(T)$. We show that $\text{REV}(D) \geq |E| \cdot (1 - 2^{-a})$, but $\text{SREV}(D) \leq 2m$ and $\text{BREV}(D) \leq 2$.

First, consider the random variable $v(M)$. We have $v(M) \leq \sum_{e=1+a}^{|E|+a} w(e)$. For any price p, in order to have $v(M) \geq p$, we must have $w(e) > 0$ for some $e \geq \log p$, as $\sum_{e=1+a}^{\log p - 1} 2^e = p - 2^{a+1} < p$. Note that there is no reason to price below 2^{1+a}. Also, by union bound, the probability that this occurs is at most $\sum_{e \geq \log p} 2^{-e} \leq 2^{1-\log p} \leq 2/p$. For any price p we could set on the grand bundle, it sells with probability at most $2/p$, so $\text{BREV} \leq 2$.

Similarly, for any price p_i, in order for the buyer to possibly be willing to purchase item i, we must have $\sum_{e \ni i} w(e) \geq p_i$. Again, in order for this to happen, we must have $w(e) > 0$ for some $e \geq \log p_i$, $e \ni i$. Again by union bound, the probability that this occurs is at most $2/p_i$. For any price p_i we could set on item i, the probability that the buyer is possibly willing to purchase item i is at most $2/p_i$, so $\text{SREV} \leq 2m$.

Consider however the following mechanism, which essentially sells the hyperedges in E separately. The mechanism allows the buyer to purchase any set/hyperedge S she chooses and charges price 2^S. Note that because we indexed the hyperedges in increasing order of size, the cheapest set that contains S is in fact S itself. By union bound ($\sum_{e=1+a}^{n+a} 2^{-e} = 2^{-a} - 2^{-n-a}$; therefore, its complement is at least $1 - 2^a$), the probability that $v \equiv 0$ is at least $1 - 2^{-a}$. Therefore, whenever $w(e) > 0$, with probability at least $1 - 2^{-a}$, the buyer will choose to purchase exactly the set e and pay 2^e. So, the revenue is at least $\sum_{e=1+a}^{|E|+a} 2^{-e} \cdot 2^e \cdot (1 - 2^{-a}) = |E| \cdot (1 - 2^{-a})$.

Finally, consider a d regular hypergraph (M, E) over m nodes with hyperedges of size k (this necessitates $d \leq \binom{m}{k}$). By definition, every node is contained in exactly d edges. Therefore, if E is the set of hyperedges used to construct D, then D has complementarity d, and $|E| = dm/k$. Taking $a \to \infty$ completes the proof.

Furthermore, we argue that this parameter correctly characterizes the degree of complementarity in our setting. Specifically, in Proposition 6, we establish extremely high lower bounds (as a function of the complementarity degree) on the approximation ratio that can be obtained by $\max\{\text{BREV}, \text{SREV}\}$ for previous measures of complementarity from the literature.

A valuation is positive hypergraph of degree at most k (PH-k) (Abraham et al. 2012) if its hypergraph representation w has only nonnegative hyperedges and only positive hyperedges S of size at most k. A valuation is Positive Supermodular of degree k (PS-k) if in its hypergraph representation, every item shares a positive hyperedge with at most k other items (and all hyperedges are nonnegative). The following proposition asserts the lower bounds for the aforementioned hierarchies.

Proposition 6. *The following hold for distributions in our settings, where hyperedge values $w(T)$ are independently drawn, and $v(S) = \sum_{T \subseteq S} w(T)$.*

1. *There exists a distribution D with only PH-k valuations in the support, for which $REV \geq \frac{1}{2m} \sum_{1 \leq i \leq k} \binom{m}{i} \max\{BREV, SREV\}$ (e.g., for PH-2, $REV \geq \Omega(m) \cdot \max\{BREV, SREV\}$).*
2. *There exists a distribution D with only PS-k valuations in the support, for which $REV \geq \frac{2^{k+1}-1}{2(k+1)} \max\{BREV, SREV\}$.*

Proof. To show Proposition 6.1, consider the distribution D given in the proof of Proposition 5, with E being the set of all hyperedges of size at most k. To show Proposition 6.2, assume for simplicity that m is divisible by $k+1$. Partition M to $m/(k+1)$ sets $M_1, M_2, \ldots M_{m/(k+1)}$, all of size $k+1$, and let E be the set of all hyperedges $S \subseteq M_i$ for all i. Every item i in M_j has neighbors only from M_j; therefore, every valuation in the support is from PS-k. The number of hyperedges is $\frac{m}{k+1} \cdot (2^{k+1} - 1)$. □

Proposition 6 also (trivially) holds for generalized hierarchies. A valuation v is MPH of degree at most k (Feige et al. 2015) if there exists a collection L of such hyperedge weight functions, so that $v(S) = \max_{\ell \in L} \{\sum_{T \subseteq S} w_\ell(T)\}$. A valuation v is maximum over PS (MPS) of degree at most k (Feldman et al. 2016) if there exists a collection L of such hyperedge weight functions, so that $v(S) = \max_{\ell \in L}\{\sum_{T \subseteq S} w_\ell(T)\}$. Because every PH-$k$ (PS-k) valuation is also trivially in MPH-k (MPS-k), Proposition 6.1 (Proposition 6.2) also holds for such MPH-k (MPS-k) valuations. In addition, consider the SM degree (Feige and Izsak 2013), which is defined as follows.

Definition 3 (Feige and Izsak 2013) (SM)**.** A valuation v is SM of degree at most k if for each item i, the number of items i' such that there exists a set $S_{i'} \not\ni i$ so that $v(S_{i'} \cup i) - v(S_{i'}) > v(S_{i'} \setminus \{i'\} \cup \{i\}) - v(S_{i'} \setminus \{i'\})$ is at most k (i.e., i''s marginal contribution to a set may *increase* by adding another item, to at most k different items).

It can be shown that PS-k $\subseteq$ SM-k. Therefore, Proposition 6.1 carries over to SM-k.

Acknowledgments

This work was done in part while S. M. Weinberg was a research fellow at Simons Institute for the Theory of Computing. This work was carried out in partial fulfillment of the requirements for the PhD degree for A. Eden and O. Friedler. This paper was accepted to the 18th ACM Conference on Economics and Computation, and a one-page abstract appeared in the proceedings.

Appendix A. Background on the Duality Framework

We first recall the duality approach of Cai et al. (2016).

Definition A.1 (Reworded from Cai et al. 2016, definitions 2 and 3)**.** A mapping $\lambda : V \times V \to \mathbb{R}^+$ is flow conserving if for all $v \in V$, $\sum_{v' \in V} \lambda(v, v') \le f(v) + \sum_{v' \in V} \lambda(v', v)$.[2] The virtual transformation associated with λ, Φ^λ, is a transformation from valuation functions in V to valuation functions in $V^\times$ (the closure of V under linear combinations) and satisfies[3]

$$\Phi^\lambda(v)(\cdot) = v(\cdot) - \frac{1}{f(v)} \sum_{v' \in V} \lambda(v', v)(v'(\cdot) - v(\cdot)).$$

In the definition, one should interpret $\lambda(\cdot,\cdot)$ as being potential Lagrangian multipliers for incentive constraints in a certain linear program to find the revenue-optimal mechanism and think of $f(v)$ flow going into each v from some super source, $\lambda(v, v')$ flow going from v to v', and all excess flow (that enters v but does not leave) as going from v to a super sink. Note that whether a given λ is flow conserving depends on the population D. Cai et al. (2016) show that Lagrangian multipliers that satisfy the flow conservation constraint yield upper bounds of the following form.

Theorem A.1 (Reworded from Cai et al. 2016, theorem 10)**.** *Let $\mathcal{M}$ be any truthful mechanism where a bidder with type v receives items $X(v)$ and pays $p(v)$. Then, for all flow-conserving λ, the expected revenue of $\mathcal{M}$ is upper bounded by its expected virtual welfare with respect to λ. That is,*

$$\mathop{\mathbb{E}}_{v \leftarrow D}\big[p(v)\big] \le \mathop{\mathbb{E}}_{v \leftarrow D}\Big[\Phi^\lambda(v)(X(v))\Big].$$

As an immediate corollary, we can obtain the following upper bound on the revenue of any truthful mechanism by observing that the bound in Theorem A.1 is maximized when $X(v)$ is deterministically $\arg\max_{S \subseteq M}\{\Phi^\lambda(v)(S)\}$.

Corollary A.1. *For all D, and all flow-conserving λ, we have*

$$REV(D) \le \mathop{\mathbb{E}}_{v \leftarrow D}\left[\max_{S \subseteq 2^M} \Phi^\lambda(v)(S)\right].$$

We begin this section by defining our flow-conserving λ and the resulting Φ^λ. Readers familiar with Cai et al. (2016) will recognize it as the natural generalization of their flow to our setting, and we will make the language as similar as possible.

We will break V into $2^m - 1$ different regions, depending on which hyperedge is the most valuable to a buyer with value v. Specifically, we say that v is in region R_A if $A = \arg\max_{T \subseteq M}\{w(T)\}$, with ties broken lexicographically. Recall that D is established by drawing w from the product distribution D', and the returned valuation v satisfies $v(S) = \max_{T \subseteq S, T \in \mathcal{C}}\{\sum_{U \subseteq T} w(U)\}$. Then, consider the following flow.

Definition A.2 (Flow for Our Benchmark)**.** If $v \in R_A$, define $w'(T) = w(T)$, for all $T \ne A$, and define $w'(A) = \min_{x > w(A)}\{x : f_A(x) > 0\}$. Set $\lambda(v', v) = \Pr_{x \leftarrow D'_A}[x \ge w'(A)] \cdot \prod_{T \ne A} f_T(w'(T)) = f(v) \cdot \frac{\Pr_{x \leftarrow D'_A}[x \ge w'(A)]}{f_A(w(A))}$ for the $v'(\cdot)$ such that $v'(S) = \max_{T \subseteq S, T \in \mathcal{C}}\{\sum_{U \subseteq T} w'(U)\}$ for all S, and $\lambda(v'', v) = 0$ for all other v''.

Proposition A.1. *The $\lambda(\cdot,\cdot)$ from Definition A.2 is flow conserving. Moreover, if $v(\cdot)$ is such that $v(S) = \max_{T \subseteq S, T \in \mathcal{C}}\{\sum_{U \subseteq T} w(U)\}$, and $v \in R_A$, then Φ^λ satisfies the following*:

$$\begin{aligned}\Phi^\lambda(v)(S) &\le \max_{T \subseteq S, T \in \mathcal{C}}\left\{\sum_{U \subseteq T, U \ne A} w(U)\right\} \\ &\quad + \max\{0, \varphi_A(w(A))\} \\ &\le \max_{T \in \mathcal{C}}\left\{\sum_{U \subseteq T, U \ne A} w(U)\right\} \\ &\quad + \max\{0, \varphi_A(w(A))\}.\end{aligned}$$

Proof. That $\lambda(\cdot,\cdot)$ is flow conserving is clear. Every $v \in R_A$ has total incoming flow of

$$f(v) \cdot \frac{\Pr_{x \leftarrow D'_A}[x \ge w(A)]}{f_A(w(A))}, \tag{A.1}$$

where $f(v)$ of this comes from the source, and the remaining $f(v) \cdot \frac{\Pr_{x \leftarrow D'_A}[x > w(A)]}{f_A(w(A))}$ comes from other types in R_A. Every $v \in R_A$ also has outgoing flow either equal to zero (if decreasing the value of $w(A)$ moves the resulting v' out of R_A) or exactly (A.1) (otherwise). In either case, the flow going out is at most the flow coming in.

Let us now compute $\Phi^\lambda(v)(S)$. Plugging into Definition A.1, we get

$$\Phi^\lambda(v)(S) = v(S) - \frac{(v'(S) - v(S)) \Pr_{x \leftarrow D'_A}[x \ge w(A)]}{f_A(w(A))}.$$

Recall that $v'(S) \ge v(S)$ for all S, and therefore, $\Phi^\lambda(v)(S) \le v(S)$ for all S. Now there are two cases to consider. In the first case, maybe $\max_{T \subseteq S, T \in \mathcal{C}}\{\sum_{U \subseteq T, U \ne A} w(U)\} = v(S)$. In other words, the set in $\mathcal{C}$ "chosen" by a consumer with valuation v does not contain A. In this case, we immediately get that $\Phi^\lambda(v)(S) \le v(S) = \max_{T \subseteq S, T \in \mathcal{C}}\{\sum_{U \subseteq T, U \ne A} w(U)\}$, as desired.

In the second case, maybe $\max_{T \subseteq S, T \in \mathcal{C}}\{\sum_{U \subseteq T, U \ne A} w(U)\} < v(S)$. In other words, the set in $\mathcal{C}$ "chosen" by a consumer with valuation v contains A. In this case, increasing $w(A)$ by any $x > 0$ increases $v(S)$ by exactly x. Therefore, we have $v'(S) = v(S) + w'(A) - w(A)$, and therefore,

$$\begin{aligned}
\Phi^{\lambda}(v)(S) &= v(S) - \frac{(w'(A) - w(A))\Pr_{x\leftarrow D'_A}[x \geq w(A)]}{f_A(w(A))} \\
&= \max_{T\subseteq S, T\in\mathcal{C}}\left\{\sum_{U\subseteq T} w(U)\right\} \\
&\quad - \frac{(w'(A) - w(A))\Pr_{x\leftarrow D'_A}[x \geq w(A)]}{f_A(w(A))} \\
&\leq \max_{T\subseteq S, T\in\mathcal{C}}\left\{\sum_{U\subseteq T, U\neq A} w(U)\right\} + w(A) \\
&\quad - \frac{(w'(A) - w(A))\Pr_{x\leftarrow D'_A}[x \geq w(A)]}{f_A(w(A))} \\
&= \max_{T\subseteq S, T\in\mathcal{C}}\left\{\sum_{U\subseteq T, U\neq A} w(U)\right\} + \varphi_A(w(A)).
\end{aligned}$$

The last line uses the definition

$$\varphi_A(w(A)) = w(A) \; - \frac{(w'(A) - w(A))\Pr_{x\leftarrow D'_A}[x \geq w(A)]}{f_A(w(A))},$$

which may seem unfamiliar to readers more familiar with virtual values for continuous distributions. Indeed, this is the right generalization of Myerson's $\varphi(\cdot)$ for continuous distributions to the discrete setting, and we refer the interested reader to section 4 of Cai et al. (2016) for more discussion. □

A.1. Ironing

The astute reader will notice that when D'_S is irregular, the bound we probably want would replace $\varphi_A(\cdot)$ with $\bar{\varphi}_A(\cdot)$. Cai et al. (2016) shows how to design a flow that accomplishes this essentially by adding cycles to λ between adjacent types to "iron out" any nonmonotonicities but for their setting of additive buyers. The exact same approach will work here. We omit a proof and refer the reader to Cai et al. (2016) for more detail. This allows us to prove Corollary 1.

Proof of Corollary 1. Simply combine Corollary A.1 and Proposition A.1, after replacing $\varphi(\cdot)$ in Proposition A.1 with $\bar{\varphi}(\cdot)$. □

Appendix B. Background on the Copies Environment

Recall that a random variable X is first-order stochastically dominated (FOSD) by random variable Y if for every x, $\Pr[X \geq x] \leq \Pr[Y \geq x]$. We remark that if X is FOSD by Y. then $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

Proof of Theorem 1 (Chawla et al. 2010)**.** Let $\mathbb{1}_q$ be an independent indicator random variable that equals one with probability q. Let $\{X_S\}_S$ be nonnegative independent random variables that are drawn from the independent distributions.

Consider a tie breaking rule among the sets, and let the event $X_S = \max_T\{X_T\}$ be true only when S also wins in the tie-breaking rule. Set $q_S = \Pr[X_S = \max_T\{X_T\}]$. So, $\sum_S q_S = 1$, set t_S s.t. $\Pr[X_S \geq t_S] = q \cdot q_S$.

Let us see that the random variable $\mathbb{1}_q \cdot X_S \cdot \mathbb{1}[X_S = \max\{X_T\}]$ is FOSD by $X_S \cdot \mathbb{1}[X_S \geq t_S]$.

For every $x \geq t_S$, it holds that $\Pr[\mathbb{1}_q \cdot X_S \cdot \mathbb{1}[X_S = \max\{X_T\}] \geq x] \leq q\Pr[X_S \geq x]$,
whereas $\Pr[X_S \cdot \mathbb{1}[X_S \geq t_S] \geq x] = \Pr[X_S \geq x]$.

For every $x < t_S$, it holds that $\Pr[\mathbb{1}_q \cdot X_S \cdot \mathbb{1}[X_S = \max\{X_T\}] \geq x] \leq q \cdot q_S$ by definition of q_S, whereas $\Pr[X_S \cdot \mathbb{1}[X_S \geq t_S] \geq x] = \Pr[X_S \geq t_S] = q \cdot q_S$ by definition of t_S. We get that

$$\begin{aligned}
\mathbb{E}\left[\max_S\{X_S\}\right] &= \sum_S \mathbb{E}[X_S \cdot \mathbb{1}[X_S = \max\{X_T\}]] \\
&= \frac{1}{q}\sum_S \mathbb{E}[\mathbb{1}_q \cdot X_S \cdot \mathbb{1}[X_S = \max\{X_T\}]] \\
&\leq \frac{1}{q}\sum_S \mathbb{E}[X_S \cdot \mathbb{1}[X_S \geq t_S]].
\end{aligned}$$

Let X_S be the random variable that first draws $x \leftarrow D'_S$ and returns $\max\{0, \varphi_S(x)\}$. Assume the distributions are regular, and refer to Chawla et al. (2010) for the irregular case. As $t_S \geq 0$, we get

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x\leftarrow D'_S}\left[\max_S\{\varphi_S(x), 0\}\right] \\
&\leq \sum_S \mathop{\mathbb{E}}_{x\leftarrow D'_S}[\varphi_S(x) \cdot \mathbb{1}[\varphi_S(x) \geq t_S]].
\end{aligned}$$

Observe that the term for each S is the expected virtual value of the mechanism that allocates to a bidder with value x if x exceeds $p_S = \inf\{x : \varphi_S(x) = t_S\}$. This allocation is achieved by posting a price p_S. By Myerson's payment identity,

$$\mathop{\mathbb{E}}_{x\leftarrow D'_S}[\varphi_S(x) \cdot \mathbb{1}[\varphi_S(x) \geq t'_S]] = \mathop{\mathbb{E}}_{x\leftarrow D'_S}[p_S \cdot \Pr[x \geq p_S]].$$

This concludes Property 1. Property 2 follows by monotonicity of φ_S (regularity of D'_S, for the irregular case refer to Chawla et al. 2010):

$$\begin{aligned}
\sum_S \Pr_{D'_S}[x \geq p_S] &= \sum_S \Pr_{D'_S}[\varphi_S(x) \geq t_S] \\
&= \sum_S q \cdot q_S = q. \quad \square
\end{aligned}$$

Appendix C. Revenue Guarantees for Pricing Bundles Require a Different Approach

Following our results, one natural question that may be raised is, instead of either pricing items or only the grand bundle, if there is a "simple" (or at least, deterministic) scheme for pricing subsets of items that gets a constant fraction of the optimal revenue. In this section, we briefly establish that a significantly new approach would be required to resolve this question. Specifically, we provide an example in which the optimal revenue from pricing subsets is a constant, but in this example, our upper bound to the optimal revenue is $O(m)$; therefore, revenue from pricing subsets cannot cover our upper bound up to a constant factor (and we would need a new upper bound to use as a starting point). This leaves the following open questions: Is there indeed an $O(m)$ gap between the optimal revenue from pricing subsets and the optimal revenue, or can one come up with a tighter upper bound to the optimal revenue that proves an $o(m)$ gap between the revenue from pricing subsets and the optimal revenue?

Example A.1. Consider m items, and a buyer without feasibility constraints (i.e., $\mathcal{C} = 2^{[m]}$), with a valuation defined by the hypergraph representation w distributed by $w([i]) = 2^{m-i}$ with probability $2^{-(m-i)}$, and $w([i]) = 0$ otherwise. The key property here is that smaller hyperedges have higher weight with lower probability.

Consider now any scheme that prices bundles. Define p_i to be the cheapest price the buyer can pay in order to purchase sets for which their union contains $[i]$. Note that p_i is monotone increasing by definition (for any $i' > i$, any union of sets that contains $[i']$ also contains $[i]$).

Let j^* be the largest index j such that $p_j \le 2 \cdot 2^{m-j}$. We prove the following two claims.

Claim A.1. The buyer will never pay more than $3 \cdot 2^{m-j^*}$.

Proof. By definition of j^*, it holds that $p_{j^*} \le 2 \cdot 2^{m-j^*}$. Therefore, the buyer can attain value from all hyperedges $[i]$ for which $i \le j^*$ at price p_{j^*}. The only hyperedges they might yet not have value for are the $[i]$ for which $i > j^*$. However, the total contribution of all such hyperedges to the buyer's value is at most $\sum_{k=j^*+1}^{m} 2^{m-k} = \sum_{k=0}^{m-(j^*+1)} 2^k < 2^{m-j^*}$.

Therefore, the buyer will always prefer the option that costs p_{j^*} to receive a superset of $[j^*]$ than even the grand bundle $[m]$ at any price $> p_{j^*} + 2^{m-j^*}$. Therefore, the buyer will never choose to purchase an option at price $> 3 \cdot 2^{m-j^*}$. □

Claim A.2. The buyer will only buy something with probability at most $2^{-(m-j^*)+1}$.

Proof. Let i^* be the smallest item for which $w([i^*]) > 0$. Similar to Claim A.1, the buyer's total value for $[m]$ (i.e., from the hyperedges $\{[k] : k > i^* - 1\}$) is strictly less than 2^{m-i^*+1}.

If $i^* > j^*$ (i.e., $i^* \ge j^* + 1$), then in order to acquire nonzero value, the buyer must purchase at least at price $p_{i^*} \ge p_{j^*+1} > 2^{m-j^*} \ge 2^{m-i^*+1}$ and result in negative utility. Therefore, if $i^* > j^*$, then the buyer will purchase nothing.

By the union bound, the probability that $i^* \le j^*$ is

$$\Pr[\exists i \le j^* : w([i]) \ne 0] \le \sum_{i=1}^{j^*} 2^{-(m-i)} \le 2^{-(m-j^*)+1}. \quad \square$$

As a conclusion of Claims A.1 and A.2, the optimal revenue from pricing bundles in our example is at most $3 \cdot 2^{m-j^*} \cdot 2^{-(m-j^*)+1} = 6$. Moreover, for our example, it is not hard to see that our benchmark (Corollary 1) is m. Indeed, recall that for the highest value in the support of a distribution D, Myerson's virtual value is simply that value. That is, because all distributions in this example have a single nonzero point mass, the nonnegative virtual value is equal to the value at all points, and therefore, our benchmark simply becomes the expected welfare (which is m).

Therefore, the optimal revenue from pricing bundles cannot approximate our benchmark, which implies that a substantially different approach is required. We conclude with a proof sketch that even randomized mechanisms can achieve revenue at best $O(1)$ in this example.

Claim A.3. No randomized mechanism can guarantee revenue > 6 for this example.

Proof (Sketch). To prove the claim, we will define a flow-conserving λ for this example D and consider the resulting Φ^λ using Corollary A.1. The proof is a "sketch" only because we do not fully expand all calculations.

Observe that each type can be completely described by an m-long bit string $\vec{t}$, with $t_i = 1$ if and only if $w([i]) > 0$. Consider the following mapping $g(\cdot)$.

- For each bit string $\vec{t}$, let $i(\vec{t}) := \min\{j, t_j = 1\}$, the smallest index such that $t_i = 1$. Observe that $i(\vec{t})$ is well defined as long as $\vec{t} \ne \vec{0}$.
- If $i(\vec{t}) = m$, or $t_{i(\vec{t})+1} = 1$, then define $g(\vec{t}) = \vec{0}$.
- Otherwise, define $g(\vec{t}) := \vec{t} + \vec{\epsilon}_{i(\vec{t})+1} - \vec{\epsilon}_{i(\vec{t})}$ (swap the coordinates $i(\vec{t})$ and $i(\vec{t}) + 1$, by adding the $(i(\vec{t}) + 1)^{st}$ standard basis vector, and subtracting the $\mathrm{i}(\vec{t})th$).

We will now define our flow-conserving λ to have $\vec{t}$ send all of its incoming flow to $g(\vec{t})$. Observe that there are no cycles in the directed graph defined by $g(\cdot)$ and that every type gets incoming flow from at most one other type. We now want to figure out just how much flow is sent into each $\vec{t}$, and then, we can compute the corresponding Φ^λ.

So, consider any type $\vec{t}$. If $t_1 = 1$, then $\vec{t}$ gets no incoming flow from anywhere. If $t_1 = 0$ (and $\vec{t} \ne \vec{0}$), then $g^{-1}(\vec{t})$ exists (and is equal to $\vec{t} - \vec{\epsilon}_{i(\vec{t})} + \vec{\epsilon}_{i(\vec{t})-1}$). Observe that the total flow incoming to $\vec{t}$ is equal to the total flow incoming to $g^{-1}(\vec{t})$, plus $f(g^{-1}(\vec{t}))$. Observe also that $f(g^{-1}(\vec{t})) = f(\vec{t}) \cdot \frac{1-2^{-(m-i(\vec{t}))}}{1-2^{-(m-i(\vec{t})+1)}}/2$. Indeed, $f(\vec{t}) = (\prod_{i,t_i=1} 2^{-(m-i)}) \cdot (\prod_{i,t_i=0} 1 - 2^{-(m-i)})$, whereas $f(g^{-1}(\vec{t})) = (\prod_{i,t_i=1} 2^{-(m-i)})(\prod_{i,t_i=0} 1 - 2^{-(m-i)}) \cdot \frac{1-2^{-(m-i(\vec{t}))}}{2^{-(m-i(\vec{t}))}} \cdot \frac{2^{-(m-i(\vec{t})+1)}}{1-2^{-(m-i(\vec{t})+1)}}$ (because we flip the bits at location $i(\vec{t})$ and $i(\vec{t} - 1)$).

We now use this to inductively compute the total flow into $\vec{t}$. Indeed, if c_i denotes the ratio of $f(g^{-1}(\vec{t}))/f(\vec{t})$ when $i(\vec{t}) = i$, and d_i is such that the flow into $\vec{t}$ is equal to $d_i f(\vec{t})$, then we have the recurrence relation: $d_{i+1} = c_i + c_i d_i$, with $c_i := \frac{1-2^{-(m-i)}}{1-2^{-(m-i+1)}}/2$, and base case $d_1 = 0$. Our goal is to solve this recurrence for d_i.

Observe that for any choice of c_i, and $d_1 = 0$, this recurrence solves to $d_{i+1} := \sum_{j=1}^{i} \prod_{\ell=j}^{i} c_\ell$. So, we just need to evaluate this sum of products for our particular definition of c_i. To compute this, observe that the product telescopes, so we have

$$\prod_{\ell=j}^{i} c_\ell = \frac{1 - 2^{-(m-i)}}{1 - 2^{-(m-j+1)}} /2^{-i+j-1}.$$

Therefore, we also get

$$\begin{aligned} d_{i+1} = \sum_{j=1}^{i} \prod_{\ell=j}^{i} c_\ell &= \sum_{j=1}^{i} \frac{1 - 2^{-(m-i)}}{1 - 2^{-(m-j+1)}} /2^{-i+j-1} \\ &\ge \sum_{j=1}^{i} 2^{-i+j-1} - 2^{-m+j-1} \\ &= 1 - 2^{-i} - \left(2^{-m+i} - 2^{-m}\right) \\ &\le 1 - 2^{-i} - 2^{-m+i}. \end{aligned}$$

This then defines the following Φ^λ:

$$\Phi^\lambda\left(\vec{t}\right)(S) = \sum_i I([i] \subseteq S) \cdot I(t_i = 1) \cdot 2^{m-i} - d_i \times \sum_i I([i] \subseteq S) \cdot I\left(g^{-1}\left(\vec{t}\right)_i = 1\right) \cdot 2^{m-i}.$$

$$= (1 - d_i) \sum_i I([i] \subseteq S) \cdot I(t_i = 1) \cdot 2^{m-i} + d_i \sum_i I([i] \subseteq S) \cdot \left(I(t_i = 1) - I\left(g^{-1}\left(\vec{t}\right)_i = 1\right)\right) \cdot 2^{m-i}$$

$$\leq (1 - d_i) \sum_i I([i] \subseteq S) \cdot I(t_i = 1) \cdot 2^{m-i}$$

$$\leq \left(2^{-i(\vec{t})+1}\right) + 2^{-m+i(\vec{t})-1}\right) \sum_i I([i] \subseteq S) \cdot I(t_i = 1) \cdot 2^{m-i}.$$

The first equality simply follows from plugging into the definition of Φ^λ. The second equality simply rearranges terms. The following inequality follows by observing that $I(t_i = 1)$ and $I(g^{-1}(\vec{t})_i = 1)$ differ only on $i(\vec{t})$ and $i(\vec{t}) - 1$. Moreover, if $[i(\vec{t})] \subseteq S$, then certainly $[i(\vec{t}) - 1] \subseteq S$ as well, so whenever the positive term is counted, the negative term is counted as well (but the negative term is larger). The final inequality follows from our lower bound on d_i. The final term is clearly maximized at $S = [m]$, so we have now established that

$$\mathrm{REV}(D) \leq \mathbb{E}_{\vec{t}}\left[\left(2^{-i(\vec{t})+1} + 2^{-m+i(\vec{t})-1}\right) \sum_i I(t_i = 1) \cdot 2^{m-i}\right].$$

So, our goal is just to upper bound the right-hand side. To compute this, consider a fixed i. The probability that $t_i = 1$ is exactly $2^{-(m-i)}$. The distribution of $i(\vec{t})$, conditioned on this, is that $i(\vec{t})$ is equal to j with probability at most $2^{-(m-j)}$ (for $j < i$), and $i(\vec{t}) = i$ with the remaining probability (it is never $> i$ because we have conditioned on $t_i = 1$). Therefore,

$$\mathbb{E}_{\vec{t}}\left[\left(2^{-i(\vec{t})+1} + 2^{-m+i(\vec{t})-1}\right) \cdot \sum_i I(t_i = 1) \cdot 2^{m-i}\right]$$

$$\leq \sum_i 2^{m-i} \cdot 2^{-(m-i)} \cdot \left(\left(1 - \sum_{j<i} 2^{-(m-j)}\right) \cdot (2^{-i+1} + 2^{-m+i-1}) + \sum_{j<i} 2^{-(m-j)} \cdot \left(2^{-(j-1)} + 2^{-m+j-1}\right)\right)$$

$$\leq \sum_i (2^{-i+1} + 2^{-m+i-1} + m2^{1-m} + 2^{-2m+2i-1}/3)$$

$$\leq 2 + 1 + 2 + 2/9 \leq 6.$$

Therefore, the optimal revenue for this example, even for randomized mechanisms, is at best six. □

Endnotes

[1] For example, when purchasing shoes, each additional pair of shoes may have a marginally decreasing value, but for any pair, the two shoes are worth more than the sum of values of each shoe by itself. For results that consider the supermodular degree, see Feige and Izsak (2013), Feldman and Izsak (2014, 2017), and Izsak (2017).

[2] This is equivalent to stating that there exists a $\lambda(v, \perp) \geq 0$ such that $\lambda(v, \perp) + \sum_{v' \in V} \lambda(v, v') = f(v) + \sum_{v' \in V} \lambda(v', v)$, which might look more similar to the wording of definition 2 in Cai et al. (2016).

[3] That is, $\Phi^\lambda(v)$ is a (possibly negative) function from 2^M to $\mathbb{R}$ and satisfies $\Phi^\lambda(v)(S) = v(S) - \frac{1}{f(v)} \sum_{v' \in V} \lambda(v', v)(v'(S) - v(S))$ for all $S \subseteq M$.

References

Abraham I, Babaioff M, Dughmi S, Roughgarden T (2012) Combinatorial auctions with restricted complements. *Proc. 13th ACM Conf. Econom. Comput.* (ACM, New York), 3–16.

Babaioff M, Immorlica N, Lucier B, Weinberg SM (2014) A simple and approximately optimal mechanism for an additive buyer. *Proc. 2014 IEEE 55th Annual Sympos. Foundations Comput. Sci. (FOCS)* (IEEE, New Brunswick, NJ), 21–30.

Bateni MH, Dehghani S, Hajiaghayi MT, Seddighin S (2015) Revenue maximization for selling multiple correlated items. *Algorithms-ESA 2015* (Springer, Patras, Greece), 95–105.

Bei X, Huang Z (2011) Bayesian incentive compatibility via fractional assignments. *Proc. 22nd Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, San Francisco), 720–733.

Briest P, Chawla S, Kleinberg R, Weinberg SM (2015) Pricing lotteries. *J. Econom. Theory* 156(6):144–174.

Cai Y, Zhao M (2017) Simple mechanisms for subadditive buyers via duality. *Proc. 49th Annual ACM SIGACT Sympos. Theory Comput., STOC 2017* (ACM, New York) 170–183.

Cai Y, Devanur NR, Weinberg SM (2016) A duality based unified approach to Bayesian mechanism design. *Proc. 48th Annual ACM SIGACT Sympos. Theory Comput.* (ACM, New York), 926–939.

Cai Y, Devanur NR, Goldner K, McAfee RP (2018) Simple and approximately optimal pricing for proportional complementarities. *Proc. 2019 ACM Conf. Econom. Comput. EC '19* (ACM, New York), 239–240.

Chawla S, Miller JB (2016) Mechanism design for subadditive agents via an ex ante relaxation. *Proc. 2016 ACM Conf. Econom. Comput. EC '16* (ACM, New York), 579–596.

Chawla S, Hartline JD, Kleinberg R (2007) Algorithmic pricing via virtual valuations. *Proc. 8th ACM Conf. Electronic Commerce* (ACM, New York), 243–251.

Chawla S, Malec D, Sivan B (2015) The power of randomness in Bayesian optimal mechanism design. *Games Econom. Behav.* 91(20): 297–317.

Chawla S, Hartline JD, Malec DL, Sivan B (2010) Multi-parameter mechanism design and sequential posted pricing. *Proc. 42nd ACM Sympos. Theory Comput.* (ACM, New York), 311–320.

Daskalakis C, Weinberg SM (2012) Symmetries and optimal multi-dimensional mechanism design. *Proc. 13th ACM Conf. Electronic Commerce* (ACM, New York), 370–387.

Daskalakis C, Deckelbaum A, Tzamos C (2013) Mechanism design via optimal transport. *Proc. 14th ACM Conf. Electronic Commerce* (ACM, New York), 269–286.

Daskalakis C, Deckelbaum A, Tzamos C (2014) The complexity of optimal mechanism design. *Proc. 25th Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Portland, OR), 1302–1318.

Day R, Milgrom P (2008) Core-Selecting package auctions. *Internat. J. Game Theory* 36:393–407.

Devanur N, Morgenstern J, Syrgkanis V, Weinberg SM (2015) Simple auctions with simple strategies. *Proc. 16th ACM Conf. Econom. Comput.* (ACM, New York), 305–322.

Dobzinski S (2007) Two randomized mechanisms for combinatorial auctions. *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques* (Springer, Princeton, NJ), 89–103.

Dobzinski S, Nisan N, Schapira M (2010) Approximation algorithms for combinatorial auctions with complement-free bidders. *Math. Oper. Res.* 35(1):1–13.

Dughmi S, Han L, Nisan N (2014) Sampling and representation complexity of revenue maximization. *Proc. Internat. Conf. Web Internet Econom.* (Springer, Beijing), 277–291.

Eden A, Feldman M, Friedler O, Talgam-Cohen I, Weinberg SM (2017) The competition complexity of auctions: A Bulow-Klemperer result for multi-dimensional bidders. *Proc. 2017 ACM Conf. Econom. Comput.* (ACM, New York), 343.

Feige U (2009) On maximizing welfare when utility functions are subadditive. *SIAM J. Comput.* 39(1):122–142.

Feige U, Izsak R (2013) Welfare maximization and the supermodular degree. *Proc. 4th Conf. Innovations Theoretical Comput. Sci.* (ACM, New York), 247–256.

Feige U, Feldman M, Immorlica N, Izsak R, Lucier B, Syrgkanis V (2015) A unifying hierarchy of valuations with complements and substitutes. *Proc. 29th AAAI Conf. Artificial Intelligence* (AAAI, Menlo Park, CA), 872–878.

Feldman M, Izsak R (2014) Constrained monotone function maximization and the supermodular degree. *Leibniz Internat. Proc. Inform. (LIPIcs), Barcelona, Spain*, vol. 28, 160–175.

Feldman M, Izsak R (2017) Building a good team: Secretary problems and the supermodular degree. *Proc. 28th Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia), 1651–1670.

Feldman M, Gravin N, Lucier B (2015) Combinatorial auctions via posted prices. *Proc. 26th Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, San Diego, CA), 123–135.

Feldman M, Friedler O, Morgenstern J, Reiner G (2016) Simple mechanisms for agents with complements. *Proc. 2016 ACM Conf. Econom. Comput., Maastricht, Netherlands* (ACM, New York), 251–267.

Feldman M, Fu H, Gravin N, Lucier B (2013) Simultaneous auctions are (almost) efficient. *Proc. 45th Annual ACM Sympos. Theory Comput.* (ACM, New York), 201–210.

Hart S, Nisan N (2013) The menu-size complexity of auctions. *Proc. 14th ACM Conf. Econom. Comput.* (ACM, New York), 565–566.

Hart S, Nisan N (2017) Approximate revenue maximization with multiple items. *J. Econom. Theory* 172(11):313–347.

Hart S, Reny PJ (2015) Maximal revenue with multiple goods: Nonmonotonicity and other observations. *Theoret. Econom.* 10(3): 893–922.

Hart S, Reny PJ (2020) The better half of selling separately. *(ACM) Trans. Econom. Comput.* 7(4):1–18.

Hartline JD, Lucier B (2015) Non-optimal mechanism design. *Amer. Econom. Rev.* 105(10):3102–3124.

Hartline JD, Kleinberg R, Malekian A (2011) Bayesian incentive compatibility via matchings. *Proc. 22nd Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, San Francisco), 734–747.

Izsak R (2017) Working together: Committee selection and the supermodular degree. *Proc. Internat. Conf. Autonomous Agents Multiagent Systems* (Springer, Sao Paulo, Brazil), 103–115.

Kleinberg R, Weinberg SM (2014) Matroid prophet inequalities and applications to multi-dimensional mechanism design. *Games Econom. Behav.* 113(7):97–115.

Lehmann D, Oćallaghan LI, Shoham Y (2002) Truth revelation in approximately efficient combinatorial auctions. *J. ACM* 49(5): 577–602.

Levin J (1997) An optimal auction for complements. *Games Econom. Behav.* 18:176–192.

Li X, Yao AC-C (2013) On revenue maximization for selling multiple independently distributed items. *Proc. Natl. Acad. Sci. USA* 110(28): 11232–11237.

Milgrom P (2007) Package auctions and exchanges. *Econometrica* 75(4):935–965.

Morgenstern J (2015) Market algorithms: Incentives, learning and privacy. PhD thesis, Stanford University, Stanford, CA.

Myerson RB (1981) Optimal auction design. *Math. Oper. Res.* 6(1):58–73.

Nguyen T, Peivandi A, Vohra R (2016) Assignment problems with complementarities. *J. Econom. Theory* 165(9):209–241.

Nisan N, Segal I (2006) The communication requirements of efficient allocations and supporting prices. *J. Econom. Theory* 129(1):192–224.

Pavlov G (2011) Optimal mechanism for selling two goods. *B. E. J. Theoretical Econom.* 11(3)1–35.

Riley J, Zeckhauser R (1983) Optimal selling strategies: When to haggle, when to hold firm. *Quart. J. Econom.* 98(2):267–289.

Rochet J-C, Chone P (1998) Ironing, sweeping, and multidimensional screening. *Econometrica* 66(4):783–826.

Rubinstein A, Weinberg SM (2015) Simple mechanisms for a subadditive buyer and applications to revenue monotonicity. *Proc. 16th ACM Conf. Econom. Comput.* (ACM, New York), 377–394.

Schechtman G (2003) Concentration, results and applications. Johnson WB, Lindenstrauss J, eds. *Handbook of the Geometry of Banach Spaces*, vol. 2 (Elsevier, Amsterdam), 1603–1634.

Thanassoulis J (2004) Haggling over substitutes. *J. Econom. Theory* 117(2):217–245.

Yao AC-C (2015) An n-to-1 bidder reduction for multi-item auctions and its applications. *Proc. 26th Annual ACM-SIAM Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Piscataway, NJ), 92–109.

Alon Eden is a postdoctoral fellow at Harvard School of Engineering and Applied Sciences. His research is focused mainly on mechanism design.

Michal Feldman is a professor of computer science at Blavatnik School of Computer Science, Tel Aviv University and a visiting scholar at Microsoft Research Herzliya. Her research focuses on the intersection of computer science, game theory, and microeconomics.

Ophir Friedler is a PhD student at Blavatnik School of Computer Science, Tel Aviv University and a marketplace optimization specialist at Outbrain Inc. His research focuses mainly on algorithmic mechanism design.

Inbal Talgam-Cohen is an assistant professor and Taub Fellow in the Computer Science Department, Technion–Israel Institute of Technology. Her interdisciplinary work spans computer science, economics, and operations research.

S. Matthew Weinberg is an assistant professor of computer science at Princeton University. His research focuses broadly on algorithmic mechanism design and algorithms under uncertainty.