

INAPPROXIMABILITY OF TRUTHFUL MECHANISMS VIA GENERALIZATIONS OF THE VAPNIK–CHERVONENKIS DIMENSION*

AMIT DANIELY[†], MICHAEL SCHAPIRA[‡], AND GAL SHAHAF[†]

Abstract. Algorithmic mechanism design (AMD) studies the delicate interplay between computational efficiency, truthfulness, and optimality. We focus on AMD’s paradigmatic problem: combinatorial auctions. We present a new generalization of the Vapnik–Chervonenkis (VC) dimension to multivalued collections of functions, which encompasses the classical VC dimension, Natarajan dimension, and Steele dimension. We present a corresponding generalization of the Sauer–Shelah lemma and harness this VC machinery to establish inapproximability results for deterministic truthful mechanisms. Our results essentially unify all inapproximability results for deterministic truthful mechanisms for combinatorial auctions to date and establish new separation gaps between truthful and nontruthful algorithms.

Key words. AGT, auctions, VC dimension

AMS subject classifications. algorithmic game theory, computational complexity, communication complexity, combinatorics.

DOI. 10.1137/15M1045971

1. Introduction. Algorithmic mechanism design (AMD) studies computational environments in which the input to the algorithm is provided by self-interested, strategic parties, e.g., combinatorial auctions—the now paradigmatic problem of AMD: m items $1, \dots, m$ are being sold to n bidders $1, \dots, n$. Each bidder i has a valuation function $v_i : 2^{[m]} \rightarrow R$, which specifies i ’s “maximum willingness to pay” for every subset (“bundle”) of items $S \subseteq [m]$. The objective is to maximize social welfare, that is, to partition the items between the bidders in such a way that the “social welfare” $\sum_i v_i(S_i)$ is maximized, where S_i is the bundle assigned to bidder i . A prominent line of research in AMD is exploring the complex interplay between three natural desiderata: (1) computational efficiency; (2) truthfulness, i.e., incentivizing bidders to reveal their actual valuations to the mechanism; and (3) approximation guarantees.

Past studies expose inherent tensions between these desiderata and establish large gaps between truthful and nontruthful mechanisms in various environments, including combinatorial auctions [11, 15, 27]—the paradigmatic setting of AMD. Yet, despite much effort along these lines, long-standing questions remain wide open, including the “holy grail” [17]: designing a *deterministic* truthful mechanism for general combinatorial auctions that matches the approximation guarantee of the best nontruthful algorithm (or prove that no such mechanism exists). Beyond the contribution of research along these lines to AMD, it also yielded ideas and insights that are of broader

*Received by the editors November 2, 2015; accepted for publication (in revised form) September 25, 2017; published electronically January 11, 2018. This paper was published in preliminary form as part of the Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, ACM, New York, 2015, pp. 401–408.
<http://www.siam.org/journals/sicomp/47-1/M104597.html>

Funding: The first author was supported, in part, by the Google Europe Fellowship in Learning Theory.

[†]Dept. of Mathematics, The Hebrew University, Jerusalem, Israel (amit.daniely@mail.huji.ac.il, gal.shahaf@mail.huji.ac.il).

[‡]School of Computer Science and Engineering, The Hebrew University, Jerusalem, Israel (schapiram@cs.huji.ac.il).

interest. One example is the interesting connection between the Vapnik–Chervonenkis (VC) dimension and optimization over partial domains, presented in [27]. Papadimitriou, Schapira, and Singer [27] showed how classical VC machinery (namely, the Sauer–Shelah lemma [28, 29]) can be used to prove inapproximability results for deterministic truthful mechanisms in the context of combinatorial public projects. In combinatorial public projects, the goal is to allocate a subset of resources to a set of agents so that the allocated resources are jointly shared by all agents. The relatively simple structure of the output in this context, a subset of resources (as opposed to, e.g., partitions in combinatorial auctions, as discussed below) allows for a fairly straightforward application of VC machinery for establishing hardness of approximation results. Subsequent work further developed this idea in the combinatorial public projects setting [27, 6] and in the combinatorial auctions setting [18, 7, 5, 22].

While combinatorial public projects naturally lend themselves to standard VC lower bounding techniques, other environments pose a more a complex challenge. [18, 7, 5, 22] showed that sophisticated adaptations of existing VC machinery, namely, the Sauer–Shelah lemma, can yield inapproximability results for combinatorial auctions in a specific setting: Vickrey–Clarke–Groves (VCG)-based mechanisms for “capped-additive valuations”. However, as combinatorial auctions deal with *partitions* of a universe (as opposed to subsets), it is not clear how standard VC machinery can, in general, be applied to other auction contexts (general truthful mechanisms, other types of valuations, communication complexity, ...).

We present a new generalization of the VC dimension that is both natural from a combinatorial perspective, and encompasses the classical VC dimension, as well as past generalizations of the VC dimension: the Natarajan dimension [23] and the Steele dimension [30]. While many previously proposed generalizations of the VC dimension [28, 29, 23, 19, 4, 1, 10, 9] are motivated by machine learning applications—proving positive results for classification—our generalization is aimed at establishing negative results in AMD. We prove a corresponding generalization of the Sauer–Shelah lemma that generalizes several known bounds [28, 29, 23, 19, 30] and apply it to obtain new inapproximability results for deterministic truthful mechanisms and also for simplifying and unifying existing inapproximability results.

1.1. Generalizing the VC dimension. Let $\mathcal{H} \subset Y^X$ be a set of functions from X to Y . Given a subset $S \subset X$, let

$$\mathcal{H}|_S := \{f : S \rightarrow Y \mid \exists h \in \mathcal{H} \text{ s.t. } h|_S = f\}$$

denote the restriction of \mathcal{H} to S , where $h|_S$ is the restriction of each function $h \in \mathcal{H}$ to S .

Recall the classic VC dimension [31].

DEFINITION 1.1 (VC-dimension). Let $\mathcal{H} \subset \{0, 1\}^X$. A subset $S \subset X$ is shattered by \mathcal{H} if $\mathcal{H}|_S = \{0, 1\}^S$. The VC dimension of \mathcal{H} , denoted $VC(\mathcal{H})$, is the maximal cardinality of a subset $S \subset X$ that is shattered by \mathcal{H} .

The prominence of the VC dimension is largely due to the Sauer–Shelah lemma [28, 29]:

LEMMA 1.2 (Sauer–Shelah). For every $\mathcal{H} \subset \{0, 1\}^X$, $|\mathcal{H}| \leq \sum_{i=1}^{VC(\mathcal{H})} \binom{|X|}{i}$.

Our generalization of the VC dimension relies on the notion of “shattering to k values”:

DEFINITION 1.3 (k -shattering). Let $\mathcal{H} \subset Y^X$ and $k \geq 2$. We say that $A \subset X$ is k -shattered by \mathcal{H} , if for every $a \in A$ there is a set $Y_a \subset Y$ of size k such that the

following holds: Every function $f : A \rightarrow Y$ that satisfies $f(a) \in Y_a$ for all $a \in A$ is a restriction of some function in \mathcal{H} . That is, $\exists h \in \mathcal{H}$ s.t. $f = h|_A$.

Putting it differently, the sets $\{Y_a\}_{a \in A}$ are of size k and satisfy the following property: For every choice of elements $\{y_a\}_{a \in A}$ with $y_a \in Y_a$, there is some $h \in \mathcal{H}$ such that $h(a) = y_a$ for all $a \in A$.

DEFINITION 1.4 (k -dimension). We define the k -dimension of \mathcal{H} , denoted $\text{Dim}_k(\mathcal{H})$, as the maximal cardinality of a k -shattered set.

This notion of (k -)shattering differs from past generalizations of shattering to nonbinary domains (see, e.g., [7, 18]). We present the following generalization of the Sauer–Shelah lemma.

THEOREM 1.5. For every $\mathcal{H} \subset Y^X$ and for every $2 \leq k \leq |Y|$,

$$|\mathcal{H}| \leq \sum_{i=0}^{\text{Dim}_k(\mathcal{H})} \binom{|X|}{i} (k-1)^{|X|-i} \binom{|Y|}{k}^i \leq |X|^{\text{Dim}_k(\mathcal{H})} |Y|^{k \text{Dim}_k(\mathcal{H})} (k-1)^{|X|}.$$

We point out that (1) taking $k = |Y| = 2$, we get the VC dimension and the Sauer–Shelah lemma [28, 29]; (2) for general Y and $k = 2$ we get Natarajan’s dimension [23] and the bound in [19], which strengthen Natarajan’s bound [23]; and (3) for $k = |Y|$, we get Steele’s dimension and bound [30]. Our proof of Theorem 1.5 (the generalized Sauer–Shelah lemma) relies on a careful application of techniques used for proving previous bounds (e.g., [30, 23, 19]) and appears in section 2.

1.2. Implications for auctions. We now provide an intuitive exposition of the connection between our dimension/bound and combinatorial auctions. An allocation, in this context, is a function $h : X \rightarrow Y$, where X is the set of items and Y is the set of bidders. We focus, for ease of exposition, on the class of VCG-based, also known as (a.k.a.) maximal-in-range (MIR) mechanisms, and on bidders with particularly simple “single-minded” valuations: Each bidder i is only interested in a single bidder-specific bundle of items T_i , and assigns a value of 1 to all sets of items containing T_i and a value of 0 to all other bundles of items. An MIR mechanism has a fixed, predetermined, “bank of allocations” of the items to the bidders, and for every input (reported valuations) outputs the best allocation in this bank. A trivial MIR mechanism is the mechanism that always allocates all items to the bidder who values them the most, i.e., the MIR mechanism whose bank of allocations consists of the n partitions of the items of the form $(\emptyset, \dots, \emptyset, [m], \emptyset, \dots, \emptyset)$. A simple argument shows that the approximation ratio of this mechanism is $\min\{n, m\}$. We prove that this naive mechanism is essentially the best VCG-based (MIR) mechanism: No VCG-based mechanism has approximation ratio $m^{1-\epsilon}$ for any constant $\epsilon > 0$ unless $\text{NP} \subseteq \text{P}/\text{poly}$.

Suppose, for a point of contradiction, that there exists a computationally efficient MIR mechanism M with bank of allocations \mathcal{H} that has approximation ratio $m^{1-\epsilon}$. Observe that the allocations in \mathcal{H} can naturally be regarded as a collection of functions from $[m]$ to $[n] \cup \{*\}$; every allocation $A \in \mathcal{H}$ is associated with function f_A that maps every item in $[m]$ to a bidder in $[n]$, or leaves the item unallocated (mapping it to $*$), as in A . We now prove (see section 3), via a subtle argument that utilizes Theorem 1.5, that there exists a large (polynomial in n) set of bidders X and a large (polynomial in m) set of items Y , such that all partitions of the items in Y between bidders in X appear in \mathcal{H} and, in this sense, \mathcal{H} shatters the pair (X, Y) . We can now conclude that, as M optimizes exactly over the allocations in \mathcal{H} , M can compute the optimal allocation of $|X|$ items to $|Y|$ single-minded bidders—an NP-hard task (this is, in fact,

TABLE 1

Inapproximability results for truthful mechanisms. The table only includes lower bounds that rely on the truthfulness of the mechanisms. All results in the table are asymptotically tight, except the $m^{\frac{1}{3}-\epsilon}$ bound for VCG-based mechanisms w.r.t. submodular valuations. Question marks (“?”) indicate that there are no separation results between truthful and nontruthful mechanisms. Minus signs (“-”) indicate that inapproximability results for truthful mechanisms are irrelevant, either because of irrelevance of the model to the class of valuations (e.g., communication complexity when valuations are succinctly described), or because lower bounds for nontruthful algorithms match the existing upper bounds for truthful mechanisms.

	Comp.		Value		Demand		Comm.	
	All	VCG	All	VCG	All	VCG	All	VCG
General	$m^{1-\epsilon}$ [new]	$m^{1-\epsilon}$ [new]	-	-	?	$m^{1-\epsilon}$ [new]	?	$m^{1-\epsilon}$ [new]
Single-minded	-	$m^{1-\epsilon}$ [new]	-	-	-	-	-	-
Submodular	$m^{\frac{1}{2}-\epsilon}$ revisit [15]	$m^{\frac{1}{2}-\epsilon}$ rev. [15, 5]	$m^{\frac{1}{2}-\epsilon}$ rev. [11]	$m^{\frac{1}{2}-\epsilon}$ rev. [11]	?	$m^{\frac{1}{3}-\epsilon}$ [new]	?	$m^{\frac{1}{3}-\epsilon}$ [new]
Capped-additive	?	$m^{\frac{1}{2}-\epsilon}$ rev. [5]	-	-	-	-	-	-
k-dup	?	$m^{1-\epsilon}$ [new]	-	-	?	$m^{1-\epsilon}$ [new]	?	$m^{1-\epsilon}$ [new]

precisely the classical set-packing problem). A formal formulation of this argument is presented in section 5.

While simple, this example illustrates the strength of our VC machinery with respect to past utilizations of VC dimension arguments in the context of combinatorial auctions: (1) \mathcal{H} shatters a polynomial number of bidders and a polynomial number of items, and (2) every partition in which *all* items in X are assigned to bidders in Y appears in \mathcal{H} . Observe that shattering a constant number of bidders (e.g., 2 bidders in [5]) is insufficient, as combinatorial auctions with single-minded bidders are, in fact, tractable for a constant number of bidders. Also, if even a single allocation of all items in X to bidders in Y does not appear in \mathcal{H} , the reduction from a combinatorial auction with $|X|$ items and $|Y|$ single-minded bidders is no longer possible (as the optimal allocation might simply not be in M 's bank).

1.3. Inapproximability results. Our inapproximability results for combinatorial auctions are summarized in Table 1. Our results are categorized in three dimensions:

- *Representation of the input.* We consider the two standard models for accessing the “input”—the valuation functions—in combinatorial auctions: (1) the computational complexity model, in which valuation functions are succinctly encoded, and mechanisms must run in time that is polynomial in the input length; and (2) the oracle model, in which valuations are treated as black boxes that can only answer a certain type of query, and complexity is measured in terms of the number of queries. Three types of queries are commonly considered: (i) value queries; (ii) demand queries; and (iii) the communication complexity model, in which oracles can answer any type of query (addressed to a single valuation).
- *General deterministic mechanisms versus VCG-based mechanisms.* The VCG scheme for designing truthful mechanisms is the only universal technique

for designing truthful deterministic mechanisms. While a naive application of VCG is often computationally intractable, more clever uses of the VCG scheme provide the best deterministic truthful approximation algorithms for combinatorial auctions to date [16, 20]. We present inapproximability results for both general (unrestricted) deterministic mechanisms and for the important subcategory of VCG-based mechanisms

- *Classes of valuation functions.* Over the past decade, much effort has been invested in bounding the approximability guarantees of truthful mechanisms for different classes of valuation functions of interest, including general valuations [20], submodular valuations [11, 15], single-minded valuations [21], capped-additive valuations [5], and more. We present inapproximability results for several well-studied classes of valuation functions.

We shall now briefly highlight some of the new results in Table 1:

Inapproximability for general deterministic mechanisms. We prove that no computationally efficient and truthful mechanism for general valuations can (asymptotically) outperform the trivial m -approximation mechanism that bundles all items together and assigns them to the bidder who values them most (in a second-price auction). Specifically, for any choice of constant $\epsilon > 0$, there exists a class of succinctly described valuation functions such that (1) a nontruthful computationally-efficient algorithm achieves an approximation ratio of m^ϵ for this class, but (2) no computationally efficient and truthful algorithm (mechanism) can obtain an $O(m^{1-\epsilon})$ -approximation unless $\text{NP} \subseteq P/\text{poly}$. Our proof of this separation gap between truthful and nontruthful algorithms relies on the “direct hardness approach” [11] combined with a careful application of the Sauer–Shelah lemma.

Results for VCG-based mechanisms. We present several new inapproximability results for VCG-based mechanisms. Our results establish, in particular, (1) that the best deterministic approximation ratio for general valuations, $\frac{m}{\sqrt{\log(m)}}$, obtained via a simple VCG mechanism [20], is essentially tight; (2) that no VCG-based mechanism can match the approximation guarantees of truthful non-VCG mechanisms in combinatorial auctions with single-minded bidders [21] and in combinatorial auctions with multiple duplicates of each item [3]. The proof of these results rely on lower bounding the Steele dimension (yet another special case of our generalization); (3) a $m^{\frac{1}{3}}$ lower bound for submodular valuations in the communication complexity model, improving upon the $m^{\frac{1}{6}}$ -inapproximability result of Dobzinski and Nisan [13], obtained via a short and elegant proof that utilizes the Natarajan dimension (another special case of our generalized VC dimension).

Unifying all inapproximability results for deterministic mechanisms. We show that essentially all inapproximability results for deterministic truthful mechanisms in the combinatorial auctions setting (to date) can be proved in our generalized VC dimension framework. Inapproximability results that fit in our framework include the inapproximability results for (1) submodular valuations in the value queries model [11]; (2) submodular valuations in the computational complexity model [15]; and (3) capped-additive valuations [5]. We believe that, in this sense, our new approach marks the borderline of the state of the art and that further progress along these lines must thus entail inherently new ideas.

1.4. Organization. We begin with a proof of Theorem 1.5 (the generalization of the Sauer–Shelah lemma) in section 2. We then discuss the connections between Theorem 1.5 and partitions in section 3. Section 4 provides background on mechanism

design and combinatorial auctions and section 5 depicts a detailed exposition of our method, followed by a variety of inapproximability results for VCG-based mechanisms with respect to several valuation classes. In section 6 we extend this technique and use it to lower bound the performance of general truthful mechanisms. We conclude with some open questions in section 7.

2. Proof of Theorem 1.5. Our proof of Theorem 1.5 (the generalized Sauer–Shelah lemma) relies on a careful application of techniques used for proving previous bounds (e.g., [30, 23, 19]). We explain the interesting connections between this VC machinery and so called “approximate-partitions” in section 3. We use the following notation.

DEFINITION 2.1. For $m \geq d \geq 0$ and $n \geq k \geq 2$ let $M_{n,k}(d, m)$ be the maximal cardinality of a class $\mathcal{H} \subset [n]^{[m]}$ with $\text{Dim}_k(\mathcal{H}) \leq d$.

Theorem 1.5 can be restated as

$$M_{n,k}(d, m) \leq \sum_{i=0}^d \binom{m}{i} (k-1)^{m-i} \binom{n}{k}^i.$$

We use the following recursive inequality.

LEMMA 2.2. For every $2 \leq k \leq n$ and $m > d > 0$ we have

$$M_{n,k}(d, m) \leq (k-1) \cdot M_{n,k}(d, m-1) + \binom{n}{k} \cdot M_{n,k}(d-1, m-1).$$

Proof. Let $\mathcal{H} \subset [n]^{[m]}$ be a class of k -dimension d with $|\mathcal{H}| = M_{n,k}(d, m)$. For every subset $Y \subset [n]$ of size k let

- $\mathcal{H}_{Y,1}$ be the class of all functions $f \in \mathcal{H}$ such that $f(m) \in Y$ and for every $y \in Y$ there is $f' \in \mathcal{H}$ such that $f|_{[m-1]} = f'|_{[m-1]}$ and $f'(m) = y$;
- $\mathcal{H}_{Y,2}$ be the class of all functions $f \in \mathcal{H}_{Y,1}$ such that $f(m) = \max(Y)$;

finally, let

- $\mathcal{H}' = \mathcal{H} \setminus (\cup_{Y \in \binom{[n]}{k}} \mathcal{H}_{Y,2})$.

Clearly,

$$|\mathcal{H}| \leq |\mathcal{H}'| + \sum_{Y \in \binom{[n]}{k}} |\mathcal{H}_{Y,2}|.$$

The proof of the lemma thus follows from the following two claims:

Claim 1: $|\mathcal{H}'| \leq M_{n,k}(d, m-1)(k-1)$. Indeed, for every function $f \in \mathcal{H}'$ there are at most $k-1$ functions $f' \in \mathcal{H}'$ with $f'|_{[m-1]} = f|_{[m-1]}$. Otherwise, there are k different $f_1, \dots, f_k \in \mathcal{H}'$ that coincide on $[m-1]$. This implies a contradiction as one of these functions belongs to $\mathcal{H}_{\{f_1(m), \dots, f_k(m)\}, 2}$. Therefore, we conclude that $|\mathcal{H}'| \leq |\mathcal{H}'_{[m-1]}| \cdot (k-1) \leq M_{n,k}(d, m-1) \cdot (k-1)$.

Claim 2: For every $Y \in \binom{[n]}{k}$, $|\mathcal{H}_{Y,2}| \leq M_{n,k}(d-1, m-1)$. Indeed, we note that $|\mathcal{H}_{Y,2}| = |\mathcal{H}_{Y,2}|_{[m-1]} = |\mathcal{H}_{Y_1}|_{[m-1]}$. We also note that $\text{Dim}_k(\mathcal{H}_{Y_1}|_{[m-1]}) = d-1$. Otherwise, we could add m to a d -elements subset of $[m-1]$ that is k -shattered to obtain a $(d+1)$ -elements subset of $[m]$ that is k -shattered. \square

The next simple lemma calculates the values of $M_{n,k}(d, m)$ for the pairs d, m for which the above recursion inequality does not apply.

LEMMA 2.3. For every $2 \leq k \leq n$ and every $m \geq 1$ we have

- $M_{n,k}(0, m) = (k-1)^m$;
- $M_{n,k}(m, m) = n^m$.

The proof of the lemma is straightforward. We will upper bound $M_{n,k}(d, m)$ in terms of the numbers $N_{n,k}(d, m)$ that are defined recursively as follows:

$$N_{n,k}(d, m) = (k-1) \cdot N_{n,k}(d, m-1) + \binom{n}{k} \cdot N_{n,k}(d-1, m-1).$$

With the initial conditions, $N_{n,k}(m, m) = n^m$ and $N_{n,k}(0, m) = (k-1)^m$. By Lemma 2.2 it is clear that $M_{n,k}(m, d) \leq N_{n,k}(m, d)$. Therefore, Theorem 1.5 follows from the following lemma.

LEMMA 2.4. *For every $m \geq d \geq 1$ and $n \geq k$ we have*

$$(1) \quad N_{n,k}(d, m) \leq \sum_{i=0}^d \binom{m}{i} (k-1)^{m-i} \binom{n}{k}^i.$$

Proof. We use induction on $m+d$. We first deal with the cases $d=0$ and $d=m$. The case $d=0$ is clear. For the case $d=m$, we split into two subcases: $n=k$ and $k < n$. If $k < n$ then the last term, in the right-hand side (r.h.s.) of (1) upper bounds $N_{d,m}$. If $k=n$ then the r.h.s. of (1) counts the number of functions in $[n]^{[d]}$, and therefore equal to $N_{n,k}(d, m) = n^d$. To see that, suppose that in order to choose a function in $[n]^{[d]}$, we first choose the preimage of 1 and, after that, we assign values in $\{2, \dots, n\}$ to the remaining points in $[d]$.

We proceed to the induction step. After dealing with the cases $d=0$ and $d=m$, we may assume that $m > d \geq 1$. Now, by the induction hypothesis and the definition of $N_{n,k}(d, m)$ we have

$$\begin{aligned} N_{n,k}(d, m) &= (k-1) \cdot N_{n,k}(d, m-1) + \binom{n}{k} \cdot N_{n,k}(d-1, m-1) \\ &\leq (k-1) \cdot \sum_{i=0}^d \binom{m-1}{i} (k-1)^{m-i-1} \binom{n}{k}^i + \binom{n}{k} \cdot \sum_{i=0}^{d-1} \binom{m-1}{i} (k-1)^{m-i-1} \binom{n}{k}^i \\ &= \sum_{i=0}^d \binom{m-1}{i} (k-1)^{m-i} \binom{n}{k}^i + \sum_{i=0}^{d-1} \binom{m-1}{i} (k-1)^{m-i-1} \binom{n}{k}^{i+1} \\ &= \binom{m-1}{0} (k-1)^m \binom{n}{k}^0 + \sum_{i=1}^d \left[\binom{m-1}{i} + \binom{m-1}{i-1} \right] (k-1)^{m-i} \binom{n}{k}^i \\ &= \binom{m}{0} (k-1)^m \binom{n}{k}^0 + \sum_{i=1}^d \binom{m}{i} (k-1)^{m-i} \binom{n}{k}^i \\ &= \sum_{i=0}^d \binom{m}{i} (k-1)^{m-i} \binom{n}{k}^i. \quad \square \end{aligned}$$

3. Shattering versus approximation. We now discuss the connections between Theorem 1.5 (the generalized Sauer–Shelah lemma) and so-called approximate-partitions. We consider collections \mathcal{H} of allocations of an item set X to a set of indices Y . We will show that if \mathcal{H} (even very loosely) in some sense approximates the collection of all partitions, then there are large subsets $S \subset X$ and $A \subset Y$ for which *all* the partitions of items in S to indices in A belong to \mathcal{H} . This property will then allow

us to embed a computationally hard problem in this setting (see section 5), and thus show intractability in a variety of scenarios.

We first define the notions of shattering by collections of allocations (and in particular we accommodate scenarios that involve duplicate elements). Then, we define two approximation notions for collections of allocations, and prove shattering results.

3.1. Shattering allocations. We extend the notion of shattering to allocations, which are more natural than functions in the context of combinatorial auctions. Let X be a set of items and Y a set of indices. An *allocation* is a pairwise disjoint collection $\{S_y\}_{y \in Y}$ of subsets of X . If $\cup_{y \in Y} S_y = X$, we say that the allocation is a partition. We denote the collection of all allocations of X to the indices Y by $P(X, Y)$. The collection of allocations naturally corresponds to the collection of functions $f : X \rightarrow Y \cup \{*\}$ (here, $f^{-1}(*)$ is the set of items that were not allocated to any index). We will freely alternate between these two representations. We say that a collection \mathcal{H} of allocations *shatters* a pair of sets $S \subset X$ and $A \subset Y$ if all partitions of S to indices in A are induced by allocations from \mathcal{H} . Namely, we have the following.

DEFINITION 3.1. *Let $\mathcal{H} \subset (Y \cup \{*\})^X$ be a collection of allocations. We say that a pair of subsets $S \subset X$ and $A \subset Y$ is shattered if $A^S \subset \mathcal{H}|_S$.*

We wish to accommodate scenarios in which each item have d (identical) copies. To this end we define the following.

DEFINITION 3.2. *Let $d \geq 1$. A d -duplicate allocation of a set X to indices Y is a collection $\{S_y\}_{y \in Y}$ of subsets of X such that every $x \in X$ belongs to $\leq d$ subsets from $\{S_y\}_{y \in Y}$.*

Note that 1-duplicate allocation is just a simple allocation. We denote by $P_d(X, Y)$ the set of all d -duplicate allocations. As with standard allocations, we say that a collection \mathcal{H} of d -duplicate allocations shatters a pair of sets $S \subset X$ and $A \subset Y$ if every partition of S to the indices A is induced by \mathcal{H} . Namely, we let $\mathcal{H}|_{S,A} := \{f : S \rightarrow A \mid \exists \{S_y\}_{y \in Y} \in \mathcal{H} \forall y \in A, f^{-1}(y) = S_y \cap S\}$, and define the following.

DEFINITION 3.3 (shattering of allocations). *Let \mathcal{H} be a collection of d -duplicate allocations of a set X to indices Y . A pair of subsets $S \subset X, A \subset Y$ is shattered by \mathcal{H} , if $A^S = \mathcal{H}|_{S,A}$.*

We will also use the notion of shattering of a single index.

DEFINITION 3.4. *Let \mathcal{H} be a collection of d -duplicate allocations of a set X to indices Y . A pair $S \subset X, a \subset Y$ is shattered by \mathcal{H} , if for every $T \subset S$ there is $f \in \mathcal{H}|_{S,Y}$ with $T = f^{-1}(a)$.*

Note that if a pair $S \subset X, A \subset Y$ is shattered then, for every $a \in A$, the pair S, a is shattered.

3.2. Approximate containment. Let X be a set of m items and Y a set of n indices.

DEFINITION 3.5. *Let $\mathcal{H} \subset P_d(X, Y)$ be a collection of allocations and let $\alpha \geq 1$. We say that \mathcal{H} has the α -containment property if for every allocation $\{T_y\}_{y \in Y}$ of X , there is an allocation $\{S_y\}_{y \in Y} \in \mathcal{H}$ such that $\frac{1}{\alpha}$ percent of the nonempty sets in $\{T_y\}_{y \in Y}$ are covered by the corresponding set in $\{S_y\}_{y \in Y}$. Namely,*

$$|\{y \mid \emptyset \neq T_y \subset S_y\}| \geq \frac{1}{\alpha} |\{y \mid \emptyset \neq T_y\}|.$$

In the context of auctions, a set of allocations with the α -containment property corresponds to a VCG-based mechanism that guarantees an α -approximation w.r.t. single-minded bidders (see section 5).

THEOREM 3.6. *Let $\epsilon > 0$, $\mathcal{H} \subset P_d(X, Y)$, and assume that $n \geq m^{1-3\epsilon/4}$ and \mathcal{H} has the $m^{1-\epsilon}$ -containment property. There exists a shattered pair $S \subset X$, $A \subset Y$ of sizes $\tilde{\Omega}(m^{3\epsilon/4})$ and $m^{\epsilon/4}$.*

Proof. For the sake of simplicity, let us restrict our attention to the case that the number of indices, n , is $m^{1-\frac{3\epsilon}{4}}$. We assume also that $m^{\frac{3\epsilon}{4}} = \frac{m}{n}$ and $k := m^{\frac{\epsilon}{4}}$ are integers.

Consider a random partition, $\{T_y\}_{y \in Y}$, of X that gives exactly $m^{\frac{3\epsilon}{4}}$ items to every index $y \in Y$. Since \mathcal{H} has the $m^{1-\epsilon}$ -containment property, at least k sets in $\{T_y\}_{y \in Y}$ must be covered by some allocation in \mathcal{H} . Namely, there is some $\{S_y\}_{y \in Y} \in \mathcal{H}$ such that $T_y \subset S_y$ for at least k indices in Y .

In addition to the random partition, consider now a subset $A \subset Y$ of k indices, chosen uniformly at random and independently from $\{T_y\}_{y \in Y}$. Since k indices must be covered (i.e., there is an allocation in \mathcal{H} such that the sets corresponding to these indices contain the sets sampled to these indices), the probability that all indices in A (over the choice of both A and $\{T_y\}_{y \in Y}$) are covered is $\geq \frac{1}{\binom{n}{k}} \geq \frac{1}{\binom{m}{k}} \geq m^{-k}$. Hence, there exists a (fixed) set A of k indices which are covered with probability (w.p.) $\geq m^{-k}$ when their corresponding sets are sampled at random (by the above distribution).

By conditioning on the set which is the union of the sets corresponding to the indices in A , it follows that there exists a set $S \subset X$ with $|S| = k \cdot m^{\frac{3\epsilon}{4}} = m^\epsilon$ such that if a function $f : S \rightarrow A$ is sampled uniformly at random among all functions with $\forall i, j \in A$, $|f^{-1}(i)| = |f^{-1}(j)|$, then w.p. $\geq m^{-k}$, it holds that $f \in \mathcal{H}_{S,A}$.

It follows that $|\mathcal{H}_{S,A}| \geq k^{m^\epsilon} \cdot m^{-k} \cdot m^{-\epsilon k}$ (the term $m^{-\epsilon k}$ is a lower bound on the probability that a uniformly chosen function from S to A will satisfy $\forall i, j \in A$, $|f^{-1}(i)| = |f^{-1}(j)|$).

By Theorem 1.5, we have that $|\mathcal{H}_{S,A}| \leq (k-1)^{m^\epsilon} m^{\epsilon \text{Dim}_k(\mathcal{H}_{S,A})}$. It follows that

$$\begin{aligned} m^{\epsilon \text{Dim}_k(\mathcal{H}_{S,A})} &\geq \left(\frac{k}{k-1}\right)^{m^\epsilon} \cdot m^{-k} \cdot m^{-\epsilon k} \\ &\geq \left(\frac{k}{k-1}\right)^{m^\epsilon} \cdot m^{-2k} \\ &= e^{\ln\left(\frac{k}{k-1}\right)m^\epsilon - 2k \ln(m)} \\ &= e^{\ln\left(1+\frac{1}{k-1}\right)m^\epsilon - 2k \ln(m)} \\ &\geq e^{\frac{1}{2k}m^\epsilon - 2k \ln(m)}. \end{aligned}$$

The last inequality is correct for large k since $\ln(1+x) = x + o(x)$. Taking logarithms we conclude that $\text{Dim}_k(\mathcal{H}_{S,A}) = \tilde{\Omega}(m^{\frac{3\epsilon}{4}})$. \square

3.3. Approximate intersection. Let X be a set of m items and Y a set of n indices.

DEFINITION 3.7. *Let $\mathcal{H} \subset P_d(X, Y)$ be a collection of allocations and let $\alpha \geq 1$. We say that \mathcal{H} has the α -intersection property if for every allocation $\{T_y\}_{y \in Y}$ of X , there is an allocation $\{S_y\}_{y \in Y} \in \mathcal{H}$ that agrees with $\{T_y\}_{y \in Y}$ on $\frac{1}{\alpha}$ percent of the items. Namely,*

$$\sum_{y \in Y} |T_y \cap S_y| \geq \frac{1}{\alpha} \sum_{y \in Y} |T_y|.$$

In the context of auctions, a set of allocations with the α -intersection property corresponds to a VCG-based mechanism that guarantees an α -approximation w.r.t. 0/1-additive valuations (see section 5).

THEOREM 3.8. *Suppose $n \geq m^{\frac{1}{k+1}-\epsilon}$ and $\mathcal{H} \subset P(X, Y)$ has the $\frac{n}{1+2k}$ -intersection property. There is a shattered pair $S \subset X$, $A \in \binom{Y}{k}$ with $|S| \geq \frac{m^{(k+1)\epsilon}}{\log_2(m)}$.*

We will use the following lemma from [7, Lemma 5].

LEMMA 3.9 (see [7]). *Assume that $\mathcal{H} \subset P(X, Y)$ has the $\frac{n}{1+2k}$ -intersection property. There exists a set $S \subset X$ of size $|S| \geq \frac{m}{n}$ such that $|\mathcal{H}_{S,Y}| \geq (1+k)\frac{m}{n}$.*

Proof of Theorem 3.8. For simplicity, we assume that $n = m^{\frac{1}{k+1}-\epsilon}$ (otherwise, we will look on the restriction of \mathcal{H} to a fixed set of $m^{\frac{1}{k+1}-\epsilon}$ indices). By Lemma 3.9, there is a subset U of size $\geq \frac{m}{n}$ such that $|\mathcal{H}_{U,Y}| \geq (k+1)\frac{m}{n}$. Applying Theorem 1.5 we get that

$$(k+1)\frac{m}{n} \leq m^{\text{Dim}_k(\mathcal{H})} n^{k \text{Dim}_k(\mathcal{H})} (k-1)\frac{m}{n} \leq m^{2 \text{Dim}_k(\mathcal{H})} (k-1)\frac{m}{n}.$$

Taking the logarithm we get a k -shattered set $T \subset U$ of size $\frac{\log_2\left(\frac{k+1}{k-1}\right)\frac{m}{n}}{2 \log_2(m)} \geq \frac{\frac{m}{n}}{2(k-1)\log_2(m)}$. Let $\{Y_a\}_{a \in T}$ be a collection of subsets of Y that indicates that T is k -shattered. Since there are at most $\binom{n}{k} \leq \frac{n^k}{(k-1)^2}$ possible options for each Y_a , by the pigeonhole principle, there is a subset $S \subset T$ of size $\frac{\frac{m}{n}}{\log_2(m)} \geq \frac{\frac{m}{n}}{\log_2(m)^{1-(k+1)\epsilon}} \geq \frac{m^{(k+1)\epsilon}}{\log_2(m)}$ such that all the subsets $\{Y_a\}_{a \in S}$ are the same and equal to some $A \in \binom{Y}{k}$. This shows that the pair (S, A) is shattered. \square

4. Background: Mechanisms and combinatorial auctions.

4.1. Mechanisms for auctions. In a *combinatorial auction*, m items are offered for sale to n bidders. Each bidder $i \in [n]$ has a private valuation function $v_i : 2^{[m]} \rightarrow [0, \infty)$ that is *normalized* (i.e., $v_i(\emptyset) = 0$) and *nondecreasing* (if $S \subset T$ then $v_i(S) \leq v_i(T)$). A *mechanism* is an algorithm that takes as input bidders' valuation functions (either directly or via oracle access) and outputs an *allocation* (i.e., a collection $S = \{S_i\}_{i \in [n]}$ of pairwise disjoint subsets of $[m]$) of items to the bidders, and a vector of prices $p_i \geq 0$ $i \in [n]$ that each bidder has to pay.

We will consider combinatorial auctions under different assumptions on the valuation functions, and also *combinatorial auction with duplicates*: d identical units of each item are available but each bidder desires at most 1 unit of each item, and the mechanism can output a d -duplicate allocation, i.e., a collection $S = \{S_i\}_{i \in [n]}$ of subsets of $[m]$ such that each $x \in [m]$ belongs to at most d of these subsets.

Important desiderata in this context are the following:

- *Good outcomes.* We seek mechanisms that provide “good” approximations to the optimum social welfare, i.e., to the allocation $S^* = \{S_i^*\}_{i \in [n]}$ that maximizes the expression $\sum_i v_i(S_i^*)$.
- *Truthfulness (a.k.a. incentive compatibility).* Under truthful mechanisms, a bidder maximizes his utility $v_i(S_i) - p_i$ by reporting his actual valuation to the mechanism regardless of other bidders' reports (that is, truthfulness is a dominant strategy for each bidder).
- *Tractability.* We will consider both the standard computational complexity model and the query complexity model (see the introduction).

DEFINITION 4.1 (MIR/VCG-based mechanism). *A mechanism M is called MIR if there exists a bank of d -duplicate allocations \mathcal{H} such that the output of M for every n -tuple of valuation function v_1, \dots, v_n is an allocation S that maximizes $\sum_{i=1}^n v_i(S_i)$ over all allocations $S \in \mathcal{H}$.*

4.2. Valuation classes and complexity models. Research on combinatorial auctions investigates the implications of different restrictions on valuation functions. We consider the following classes of valuations:

- *General*: all (nonnegative, nondecreasing, and normalized) valuations.
- *Multiminded* (0/1): all valuations $v : 2^{[m]} \rightarrow \{0, 1\}$.
- *Single minded* (0/1): all valuations of the form $v(S) = 1[T \subset S]$ for some $T \subset [m]$.
- *Subadditive*: all valuations satisfying $v(S \dot{\cup} T) \leq v(S) + v(T)$.
- *XOS*: all valuations of the form $v(S) = \max_j w_j^T \chi(S)$, where each w_j is a vector in \mathbb{R}_+^n (here, $\chi(S)$ is the characteristic vector of S).
- *Submodular*: all valuations satisfying $v(S \cup T) \leq v(S) + v(T) - v(S \cap T)$.
- *Additive*: all valuations of the form $v(S) = w^T \chi(S)$, where $w \in \mathbb{R}_+^n$.
- *0/1-Additive*: all valuations of the form $v(S) = w^T \chi(S)$, where $w \in \{0, 1\}^n$.

These classes form the following hierarchies [25]:

$$0/1\text{-Additive} \subset \text{Additive} \subset \text{Submodular} \subset \text{XOS} \subset \text{Subadditive} \subset \text{General}$$

and

$$\text{Single minded} \subset \text{Multiminded} \subset \text{General} .$$

All inclusions above are strict.

A naive representation of the “input” to combinatorial auctions—the valuation functions—requires specifying 2^m numbers. In contrast, mechanisms are expected to be efficient in the natural parameters of the problem: the number of bidders n and the number of items m . We consider the two standard models for accessing the in combinatorial auctions:

- *The computational complexity model*, in which valuation functions are succinctly encoded and mechanisms must run in time that is polynomial in the input length; for single-minded bidders, for instance, the encoding of valuation function v_i is simply the bundle of items bidder i desires.
- *The oracle model*, in which valuations are treated as black boxes that can only answer certain types of queries, and complexity is measured in terms of the number of queries. Three types of queries are commonly considered: (i) a *value query* to valuation v_i is a subset $S \subseteq [m]$ and the reply is the value $v_i(S)$; (ii) a *demand query* to v_i is a vector $p = (p_1, \dots, p_m)$ of “item prices,” and the reply is bidder i ’s “demand” at these prices, that is, a subset S that maximizes the expression $v_i(S) \sum_{j \in S} p_j$; and (iii) the *communication complexity model*, in which providing a unifying framework for essentially all existing inapproximability results for combinatorial auctions to date, oracles can answer any kind of query (addressed to a single valuation) and the complexity of the mechanism is measured in terms of the number of bits transmitted.

5. Inapproximability for VCG-based mechanisms. We now show how our VC dimension arguments can be applied to lower bound the approximability of VCG-based mechanisms. Consider the classical combinatorial auction setting: m items $1, \dots, m$, are sold to n bidders $1, \dots, n$, and each bidder i has a private valuation

function over bundles of items $v_i : 2^{[m]} \rightarrow R$. The objective is to partition the items between the bidders so as to maximize $\sum_i v_i(S_i)$, where S_i is the bundle assigned to bidder i . A more general setting is that of a combinatorial auction with duplicates (studied, e.g., at [3], [14], [8]), in which d identical units of each item are available but each bidder desires at most 1 unit of each item. The mechanism can now output a d -duplicate allocation, i.e., a collection $S = \{S_i\}_{i \in [n]}$ of subsets of $[m]$ such that each $x \in [m]$ belongs to at most d of these subsets.

A VCG-based, a.k.a. MIR, mechanism has a fixed *bank* of allocations $\mathcal{H} \subset P_d([m], [n])$ such that the output of M for every n -tuple of valuation function v_1, \dots, v_n is an allocation S that maximizes $\sum_{i=1}^n v_i(S_i)$ over all allocations $S \in \mathcal{H}$. Charging VCG prices from the bidders ensures that such mechanisms are truthful.

We lower bound the approximability of MIR mechanisms in two steps:

Step I: Shattering a large number of bidders and items (the VC step). Using the machinery developed in the previous section we show that a “good” approximation ratio implies shattering of a large number of bidders and items by \mathcal{H} . Consider the following two of the simplest classes of valuation functions:

- *Single minded (0/1):* all valuations of the form $v(S) = 1[T \subset S]$ for some $T \subset [m]$.
- *Additive (0/1):* all valuations of the form $v(S) = w^T \chi(S)$, where $w \in \{0, 1\}^n$.

Intuitively, these valuations correspond to the containment and the intersection properties defined in the previous section. Shattering is thus immediately guaranteed by the following corollaries of Theorem 3.6 and Theorem 3.8, respectively.

COROLLARY 5.1. *Let $\epsilon > 0$, and let $\mathcal{H} \subset P_d([m], [n])$ such that $n \geq m^{1-3\epsilon/4}$ and \mathcal{H} has an approximation ratio of $m^{1-\epsilon}$ w.r.t. single-minded valuations. Then, there exist subsets $S \subset [m]$ and $A \subset [n]$ of sizes $\Omega(m^{3\epsilon/4})$ and $m^{\epsilon/4}$ that are shattered by \mathcal{H} .*

COROLLARY 5.2. *Suppose $n \geq m^{\frac{1}{k+1}-\epsilon}$ and $\mathcal{H} \subset P([m], [n])$ has an approximation ratio of $m^{\frac{1}{k+1}-\epsilon}$ w.r.t. 0/1-additive valuations. Then, there exists a shattered pair $S \subset [m]$, $A \in \binom{[n]}{k}$ with $|S| \geq \frac{m^{\frac{(k+1)\epsilon}{\log_2(m)}}}{\log_2(m)}$.*

Clearly, these corollaries assert that if mechanism M obtains a suitable approximation ratio w.r.t. any collection of valuations \mathcal{F} that contains one of these two basic valuation classes, the (appropriate) shattering bound holds.

Step II: Reducing from a hard problem. As the mechanisms considered are MIR, the mechanism can optimize exactly over the shattered set of bidders and items—an intractable task for many classes of valuations. We leverage this to prove both computational complexity and communication complexity results. We emphasize that our method only shows that a shattered set exists, but does not provide an efficient algorithm to find it. Hence, the aforementioned efficient mechanism is nonuniform. Namely, it assumes that the identity of the shattered set is given. As a result, our underlying hardness assumption is that $\text{NP} \not\subseteq \text{P/poly}$, and not that $\text{NP} \neq \text{P}$.

The remainder of this section illustrates how we can adapt this idea and apply it in specific contexts and with respect to several classes of valuation functions. The following notion of *Approximation w.r.t. Valuations Classes* will thus be a useful one:

DEFINITION 5.3. *Fix a collection of valuations \mathcal{F} . Let $\mathcal{H} \subset P_d(X, Y)$ be a collection of allocations. We say that \mathcal{H} has an approximation ratio of $\alpha \geq 1$ w.r.t. \mathcal{F} if for every choice of valuations $\{v_y\}_{y \in Y} \in Y \subset \mathcal{F}$ and every partition $\{S_y\}_{y \in Y} \in P(X, Y)$*

there is an allocation $\{T_y\}_{y \in Y} \in \mathcal{H}$ such that

$$\sum_{y \in Y} v_y(T_y) \geq \frac{1}{\alpha} \cdot \sum_{y \in Y} v_y(S_y).$$

It is easy to see that \mathcal{H} has an approximation ratio of α w.r.t. \mathcal{F} if and only if running an MIR mechanism using \mathcal{H} as the bank of allocations has an approximation ratio of α w.r.t. \mathcal{F} . We will consider questions of the following type.

QUESTION 5.4. *Given a collection \mathcal{F} of valuation, what is the best (i.e., smallest) approximation ratio $\alpha \geq 1$ for which there exists a collection of allocations with approximation ratio of α and no large shattered pair of sets?*

5.1. Single- and multi-minded valuations. Recall that single-minded bidders are interested in a single bundle of items and are willing to pay a fixed price for the whole bundle (or for any superset), and pay nothing in any other case. Though simply defined and easily represented, these valuation are complex enough such that even without requiring a mechanism to be incentive compatible, one could not expect an approximation ratio better than $O(\sqrt{m})$. On the other hand, one of the few positive results in this scenario is a greedy algorithm [21] that obtains an $O(\sqrt{m})$ -approximation ratio while maintaining incentive compatibility. One can verify that this greedy mechanism is not VCG-based, but still expect a similar performance from VCG-based mechanisms. Our first theorem asserts that this is not the case.

5.1.1. Single-minded computational bound.

THEOREM 5.5. *For every $\epsilon > 0$, no efficient VCG-based mechanism obtains an approximation ratio of $m^{1-\epsilon}$ w.r.t. single-minded valuations unless $NP \subseteq P/poly$.*

In cases where there are d duplicate copies of each item, one might hope to boost up the poor performance of mechanisms, predicated by the above theorem. Indeed, a mechanism devised by Bartal, Gonen, and Nisan [3], is both incentive compatible, have an approximation ratio of $dm^{\frac{1}{d-2}}$, and is efficient whenever demand queries can be carried out efficiently (in particular, it is efficient for single-minded bidders).

Once again, one might hope that the universal technique of VCG-based mechanisms can achieve similar (or even better) approximation ratios. A generalization of Theorem 5.5 shows that it is impossible.

THEOREM 5.6. *Assuming the unique games conjecture (UGC), for every constant d and every $\epsilon > 0$, no efficient VCG-based mechanism with d duplicates obtains an approximation ratio of $m^{1-\epsilon}$ w.r.t. single-minded valuations unless $NP \subseteq P/poly$.*

Proof. Suppose that M is VCG based, and obtains an approximation ratio of $m^{1-\epsilon}$ for combinatorial auctions with d duplicates w.r.t. single-minded valuations. Let $\mathcal{H} \subset P_d([m], [n])$ denote its bank of allocations, and let $S \in P([m], [n])$ represent single-minded bidders (bidder i desires the bundle S_i). The approximation ratio implies there must be some $T \in \mathcal{H}$ such that

$$n \leq m^{1-\epsilon} \cdot |\{i | S_i \subset T_i\}|.$$

Therefore, by Corollary 5.1, there is a shattered pair $B \subset [n]$, $X \subset [m]$ of size $m^{\epsilon/4}$, $\Omega(m^{3\epsilon/4})$, respectively.

We now use the existence of the shattered pair in order to show computational hardness, using a reduction from the k -packing promise problem. In this problem,

the input consists of r sets $S_1, \dots, S_r \subset U$ and an integer $C > 0$, and the goal is to distinguish between the following two options:

- *Positive*: There are C sets in S_1, \dots, S_n that are pairwise disjoint.
- *Negative*: For every collection \mathcal{F} of C sets out of S_1, \dots, S_n , there is some $x \in U$ that is covered by k sets of \mathcal{F} .

Observe that setting $k = 2$ yields the famous set-packing problem, which is known to be NP-hard. We shall make use of the fact that assuming GC, this problem is NP-hard for every constant k . (The proof, based on [2], can be found in the appendix.)

The reduction is given by taking $k = d + 1$, and using the shattered X set as the universe U : Given $S_1, \dots, S_{m^{\epsilon/4}} \subset X$ and an integer $C > 0$, suppose that each bidder $i \in B$ desires S_i , and all bidders not in B have the zero valuation. Now, in the “yes” cases, since there is a collection of C pairwise disjoint subsets from $S_1, \dots, S_{m^{\epsilon/4}}$, and since the pair X, B is shattered and M is MIR, the social welfare, $\sum_i v_i(S_i)$, will be at least C . On the other hand, in the “no” cases, it is impossible to satisfy C bidders by any d -duplicate allocation, and therefore, the social welfare will be less than C . \square

5.1.2. Multiminded communication bound. Single-minded valuations are thus computationally hard to deal with, while trivial from a communicational perspective—only $O(m)$ bits are required to encode each valuation. Considering not one, but several bundles of items, changes the picture.

THEOREM 5.7. *For every constant d and $\epsilon > 0$, if a VCG-based mechanism for combinatorial auctions with d -duplicates obtains an approximation ratio of $m^{1-\epsilon}$ w.r.t. multiminded valuations, it must use exponential communication.*

While proving this theorem, we shall make use of the following definition.

DEFINITION 5.8. *Let $\mathcal{F} = \{A^s\}_{s=1}^t$ be a collection of partitions such that each A^s partition the universe X into k sets. We say that \mathcal{F} has the k -wise intersection property if for every choice of k indices $1 \leq s_1, \dots, s_k \leq t$, we have $\cap_{i=1}^k A_i^{s_i} = \emptyset$ if and only if $s_1 = \dots = s_k$.*

The following lemma asserts that a large collection of partitions with such an intersection property exists.

LEMMA 5.9 (essentially [14]). *There exists a collection of partitions $\mathcal{F} = \{A^s\}_{s=1}^t$ of the set X into k sets, such that $|\mathcal{F}| = t = e^{\frac{|X|}{k^{k+1}}}$, and \mathcal{F} having the k -wise intersection property holds.*

Proof. We use a probabilistic construction: Choose each partition A^s independently at random (that is, each element in X will be placed in A_j^s with equal probability). Fix k indices $1 \leq s_1, \dots, s_k \leq t$ at least two of which are different. We have that

$$\Pr \left[\cap_{i=1}^k A_i^{s_i} = \emptyset \right] = \prod_{j=1}^{|X|} \Pr \left[j \notin \cap_{i=1}^k A_i^{s_i} \right] \leq \left(1 - \frac{1}{k^k} \right)^{|X|} \leq e^{-\frac{|X|}{k^k}}.$$

As there are at most t^k choices of indices, using the union bound we conclude that as long as $t^k < e^{\frac{|X|}{k^k}}$ such a collection of partitions exists. \square

Proof of Theorem 5.7. Suppose that M is VCG based, and obtains an approximation ratio of $m^{1-\epsilon}$ for combinatorial auctions with d duplicates w.r.t. multiminded valuations. Let \mathcal{H} denote its bank of allocations. As single-minded valuations contained in this class, \mathcal{H} shatters a pair of subsets $B \subset [n]$, $X \subset [m]$ of sizes $d + 1$, $\Omega(m^{\epsilon/2})$.

Applying Lemma 5.9, yields a collection of partitions $\mathcal{F} = \{D^s\}_{s=1}^t \subset P(X, B)$ of size $t = e^{m^{\frac{1}{2}-\epsilon}/(d+1)^{d+2}}$ which satisfy the $(d+1)$ -wise intersection property.

We now use this collection of partitions in order to show a reduction from the *approximate disjointness problem*: This problem deals with r players, each of which holds a t -bit string that specifies a subset $A_i \subseteq \{1, \dots, t\}$. The players are required to distinguish between the following two extreme cases:

- *Positive*: $\cap_{i=1}^r A_i \neq \emptyset$.
- *Negative*: $A_i \cap A_j = \emptyset$ for all $1 \leq i \neq j \leq r$.

The communication complexity of the approximate disjointness problem is known to be $\Omega\left(\frac{t}{r^4}\right)$ [24]. The reduction is given as follows: Suppose that each of the bidders in B holds a subset $A_i \subset \{1, \dots, t\}$ which determines the multiminded valuation

$$v_i(T) = \begin{cases} 1, & \exists s \in A_i \text{ s.t. } D_i^s \subseteq T, \\ 0, & \text{else.} \end{cases}$$

Observe that if $\cap_{i=1}^r A_i \neq \emptyset$, indicated by some $s \in \cap_{i=1}^r A_i$, then the allocation A^s satisfies all bidders in B , and since the pair B, X is shattered, the social welfare will be $d+1$.

On the other hand, if $A_i \cap A_j = \emptyset$ for all $1 \leq i \neq j \leq r$, then it is not hard to see that it is impossible to satisfy all the bidders in B by *any* d -duplicate allocation, and therefore, the social welfare will be $\leq k$.

We conclude that simulating these valuations (and zero valuations for all bidders not in B) solves the approximate disjointness problem, and the communication complexity of M is therefore

$$\Omega\left(\frac{t}{(d+1)^4}\right) = \Omega\left(\frac{e^{m^{\frac{1}{2}-\epsilon}/(d+1)^{d+2}}}{(d+1)^4}\right)$$

which proves the claim. □

5.2. 0/1-additive, submodular, XOS, and subadditive valuations. Fix a set of items, X , of size m and a set of bidders, Y , of size n . In this section we consider question 5.4 for additive, submodular, XOS and subadditive valuations. Recall that

$$0/1\text{-Additive} \subset \text{Submodular} \subset \text{XOS} \subset \text{Subadditive}$$

and all the inclusions are strict. In [16] it is shown that there is a class¹ $\mathcal{H} \subset P(X, Y)$ with an approximation ratio of $m^{\frac{1}{2}}$ (even $\min(m^{\frac{1}{2}}, n)$) w.r.t. subadditive valuations and with no shattered pair $S \subset X, a \in Y$ with $|S| \geq 3$. Using this class [16] provided truthful mechanism with approximation ratio of $m^{\frac{1}{2}}$ w.r.t. subadditive valuations. As shown in [7], if we require a bit more, namely, an approximation ratio of $m^{\frac{1}{2}-\epsilon}$ with respect to just to 0/1-additive valuations, we can find a shattered pair $S \subset X, a \in Y$ for a large set S .

As we show next, if we require a stronger approximation ratio, of $m^{\frac{1}{3}-\epsilon}$, we can find a shattered pair $S \subset X, A \subset Y$ with $|A| = 2$ and a large set S . This will be useful later to establish communication lower bounds.

THEOREM 5.10. *Suppose $n \geq m^{\frac{1}{k+1}-\epsilon}$ and $\mathcal{H} \subset P(X, Y)$ has an approximation ratio of $m^{\frac{1}{k+1}-\epsilon}$ w.r.t. 0/1-additive valuations. There is a shattered pair $S \subset X, A \in \binom{Y}{k}$ with $|S| \geq \frac{m^{(k+1)\epsilon}}{\log_2(m)}$.*

¹This class is simply the collection of all allocations that either gives all items to a single bidder or, gives at most a single item to each bidder.

The theorem follows immediately from Theorem 3.8, as we note that \mathcal{H} has the α -intersection property if and only if \mathcal{H} has an approximation ratio of $m^{1-\epsilon}$ w.r.t. single 0/1-additive valuations. Theorem 3.8 shows that if a class has a good approximation ratio w.r.t. 0/1-additive, we can find a set of constantly many bidders A and a large subset $S \subset X$ such that the pair S, A is shattered. The following theorem shows that for an arbitrarily good approximation ratio $\alpha > 1$, there are classes with an approximation ratio of α w.r.t. XOS valuations with no shattered pair S, A , where S is large (say, superlogarithmic) and $|A| = \omega(\log(n))$. This indicates on limitation of the VC technique for establishing a computational lower bound on combinatorial auctions with XOS valuations.

THEOREM 5.11. *For every $l \leq |Y|$ there is a class $\mathcal{H} \subset Y^X$ such that*

- \mathcal{H} has an approximation ratio of $\frac{l}{l-1}$ w.r.t. XOS valuations;
- for $l = 2$, \mathcal{H} has an approximation ratio of 2 w.r.t. subadditive valuations;
- there is no shattered pair $S \subset X$ and $A \subset Y$ with $|S| \geq 8(l - 1) \ln(n)$ and $|A| > 24(l - 1) \ln(m)$.

Proof. For simplicity, we assume that l divides $|Y|$. Fix $l \geq 2$ and let $\mathcal{Y} = \{Y_{x,j}\}_{x \in X, 1 \leq j \leq l}$ be nonempty subsets of Y such that for every $x \in X$, $Y = Y_{x,1} \cup \dots \cup Y_{x,l}$. Define $\mathcal{H}^{\mathcal{Y}}$ to be the class of all functions $f : X \rightarrow Y$ for which there is $1 \leq i \leq l$ such that $\forall x, f(x) \notin Y_i$. As the following claim shows, for every \mathcal{Y} , $\mathcal{H}^{\mathcal{Y}}$ satisfies the first two conclusions of the theorem. Later on, we will show that for some \mathcal{Y} , $\mathcal{H}^{\mathcal{Y}}$ also satisfies the last conclusion.

CLAIM 1. *The following conditions hold:*

- $\mathcal{H}^{\mathcal{Y}}$ has an approximation ratio of 2 w.r.t. subadditive valuations.
- $\mathcal{H}^{\mathcal{Y}}$ has an approximation ratio of $\frac{l}{l-1}$ w.r.t. XOS valuations.

Proof. We prove the second part. The proof of the first part is similar. Let $\{v_y\}_{y \in Y}$ be a collection of XOS valuations and let $\{S_y\}_{y \in Y} \in P(X, Y)$. We must show that there is $\{T_y\}_{y \in Y} \in \mathcal{H}^{\mathcal{Y}}$ with

$$\sum_{y \in Y} v_y(T_y) \geq \frac{l-1}{l} \sum_{y \in Y} v_y(S_y).$$

For every $1 \leq i \leq l$ let $\{S_y^i\}_{y \in Y}$ be the allocation defined by

$$S_y^i = S_y \setminus \{x \in X \mid y \in Y_{x,i}\}.$$

The allocation $\{S_y^i\}_{y \in Y}$ is contained in one of the allocations in $\mathcal{H}^{\mathcal{Y}}$. To see that, let $f : X \rightarrow Y$ be any function for which $f|_{S_y^i}$ is the constant function y and $f(x) \notin Y_{x,i}$ for every $x \in \cup_{y \in Y} \{x' \in S_y \mid y \in Y_{x',i}\}$. It is not hard to check that $f \in \mathcal{H}^{\mathcal{Y}}$ and for all $y \in Y$, $S_y^i \subset f^{-1}(y)$.

Therefore it is enough to show that for some i ,

$$\sum_{y \in Y} v_y(S_y^i) \geq \frac{l-1}{l} \sum_{y \in Y} v_y(S_y).$$

Indeed, we will show that for random i , the expected value of the left-hand side is at least the r.h.s. To this end, it is enough to show that for any fixed $y \in Y$,

$$\frac{1}{l} \sum_{i=1}^l v_y(S_y^i) \geq \frac{l-1}{l} v_y(S_y).$$

Suppose that $v_y(U) = \max_j w_j^T \chi(U)$. Let j^* be the index for which $v_y(S_y) = w_{j^*}^T \chi(S_y)$. Now, we have

$$\sum_{i=1}^l v_y(S_y^i) \geq \sum_{i=1}^l w_{j^*}^T \chi(S_y^i) = (l-1) \cdot w_{j^*}^T \chi(S_y) = (l-1)v_y(S_y),$$

where the left equality follows from the fact that each element $x \in S_y$ appears in exactly $l-1$ many S_y^i 's (namely, all i 's except the one for which $y \in Y_{x,i}$). \square

Suppose now that for every $x \in X$, the partition $\{Y_{x,i}\}_{i \in [l]}$ is chosen at random, uniformly from all partitions of Y into l nonempty subsets, each of which is of size $\frac{|Y|}{l}$. Now, fix subsets $S \subset X$ and $A \subset Y$.

CLAIM 2. *The probability that the pair S, A is shattered is $\leq |S|^l l^{|S|} (\frac{l-1}{l})^{|A|(|S|-l)}$.*

Proof. We say that $x \in S$ is a *bad point* if for every $1 \leq i \leq l$, $A \cap Y_{x,i} \neq \emptyset$. We note that if there are l bad points, then the pair S, A is not shattered. Indeed, if x_1, \dots, x_l are bad points, then we can choose for every $1 \leq i \leq l$ some $y_{x_i} \in A \cap Y_{x_i,i}$. It is not hard to see that there is no $f \in \mathcal{H}^{\mathcal{Y}}$ satisfying $\forall i, f(x_i) = y_{x_i}$. Therefore, the pair S, A is not shattered.

It is therefore enough to bound the probability that there are less than l bad points. Indeed, for every $x \in S$, the probability that $A \cap Y_{x,i} = \emptyset$ is $\leq (\frac{l-1}{l})^{|A|}$. Hence, the probability that x is good is $\leq l(\frac{l-1}{l})^{|A|}$. It follows that the probability that there are less than l bad points is

$$\leq \sum_{i=1}^{l-1} \binom{|S|}{i} l^{|S|-i} \left(\frac{l-1}{l}\right)^{|A|(|S|-i)} \leq |S|^l l^{|S|} \left(\frac{l-1}{l}\right)^{|A|(|S|-l)}. \quad \square$$

By the claim, it follows that the probability that some pair S, A with $|S| = m' > 8(l-1)\ln(n)$ and $|A| = n' > 24(l-1)\ln(m)$ is shattered is

$$\leq \binom{n}{n'} \binom{m}{m'} m'^l l^{m'} \left(\frac{l-1}{l}\right)^{n'(m'-l)} \leq n^{n'} m^{2m'+l} \left(\frac{l-1}{l}\right)^{n'(m'-l)}.$$

Taking \ln , and using the facts that $2l \leq m'$ and that for every $0 \leq x \leq 1$ we have that $\ln(1+x) \geq \ln(2)x \geq \frac{x}{2}$, we conclude that the \ln of the last probability is

$$\begin{aligned} & n' \ln(n) + (2m' + l) \ln(m) - \ln\left(\frac{l}{l-1}\right) n'(m' - l) \\ & \leq n' \ln(n) + (2m' + l) \ln(m) - \frac{1}{2(l-1)} n'(m' - l) \\ & \leq n' \ln(n) + 3m' \ln(m) - \frac{1}{4(l-1)} n' m' \\ & = n' \left(\ln(n) - \frac{m'}{8(l-1)} \right) + m' \left(3 \ln(m) - \frac{n'}{8(l-1)} \right) < 0. \end{aligned}$$

Therefore, there is a positive probability that no such pair is shattered. Hence, there is some collection of partitions, \mathcal{Y} , for which $\mathcal{H}^{\mathcal{Y}}$ satisfies the third conclusion of the theorem. \square

In the following subsections we establish four lower bounds for submodular valuations and subclasses of submodular valuations. These demonstrate the power and

generality of the VC technique, as we show how to prove under this framework three known results and one that strengthens a previous result of [13]. Besides unification, the proofs depicted here are somewhat simpler. In particular, Theorem 5.12 substantially simplifies (and strengthens) [13].

5.2.1. Communication lower bounds on VCG-based mechanisms.

THEOREM 5.12. *For every $\epsilon > 0$, if a VCG-based mechanism has an approximation ratio of $m^{\frac{1}{3}-\epsilon}$ w.r.t. submodular valuations, it must use exponential communication.*

This result strengthens a similar result due to [13], which showed VCG mechanisms with an approximation ratio of $m^{\frac{1}{6}-\epsilon}$ w.r.t. submodular valuations must use exponential communication. In addition, the following proof is substantially simpler.

Proof. Let M be a VCG-based mechanism having an approximation ratio of $m^{\frac{1}{3}-\epsilon}$ w.r.t. submodular valuations. Let $\mathcal{H} \subset P([m], [n])$ be its bank of allocations. Using Theorem 5.10, we can find a shattered pair S, A , where $|S| \geq m^\epsilon$ and $|A| = 2$. Now, assume that all the players besides the players in $A := \{1, 2\}$ have the zero valuation, and the valuations of the bidders in A depend only on the items in S .

The output of the mechanism induces an allocation of $S = A_1 \dot{\cup} A_2$ that maximizes $v_1(A_1) + v_2(A_2)$. Therefore, the problem of finding an allocation that maximizes the sum of two submodular functions can be solved using this mechanism. Since the solution of this problem requires exponential communication [26], the theorem follows. \square

5.2.2. Computational and value queries lower bounds.

THEOREM 5.13. *For any $\epsilon > 0$, unless $NP \subseteq P/poly$, no efficient truthful mechanism has an approximation ratio of $m^{\frac{1}{2}-\epsilon}$ w.r.t. submodular valuations.*

THEOREM 5.14. *For any $\epsilon > 0$ any truthful mechanism that communicates with the bidders only by value queries, and has an approximation ratio of $m^{\frac{1}{2}-\epsilon}$ w.r.t. submodular valuations, must use exponentially many value queries.*

The first theorem was proved by [15] under the weaker assumption that $NP \neq P$. The second theorem was proved by [11]. The proof depicted here is close to the proofs in [11, 15], but the step of embedding a hard problem given a “large menu” is somewhat simplified.

Proof of Theorems 5.13 and 5.14. We will use the terminology of section 6. Assume $n = m^{\frac{1}{2}-\epsilon}$. Let M be a truthful mechanism with approximation ratio $m^{\frac{1}{2}-\epsilon}$ w.r.t. submodular valuations. By [15, Lemma 2.3], there is $1 \leq k \leq m$, $0 \leq p \leq m$ and additive valuations v_{-i} for which $|\mathcal{S}(v_{-i}, k, p)| \geq \frac{\epsilon m^{2\epsilon}}{10m^8}$. Denote $\mathcal{S} = \mathcal{S}(v_{-i}, k, p)$. By the Sauer–Shelah lemma, \mathcal{S} shatters a set $U \subset [m]$ of size $\geq m^\epsilon$ (for large enough m). Given collection B of subsets of S , consider the following valuation,

$$v_i(S) = \begin{cases} t|S|, & |S| < k, \\ tk, & |S| = k, S \in \mathcal{S} \text{ and } S \cap U \in B, \\ tk - \frac{1}{m^4}, & |S| = k, S \in \mathcal{S} \text{ and } S \cap U \notin B, \\ t(k - 2^{-k}), & |S| = k \text{ and } S \notin \mathcal{S}, \\ t(k - 2^{-|S|}), & |S| > k \text{ and } p_{v_{-i}}(S) < p + \frac{1}{m^5}, \\ tk, & |S| > k \text{ and } p_{v_{-i}}(S) \geq p + \frac{1}{m^5}. \end{cases}$$

It is not hard to check that v_i is nondecreasing, normalized, and submodular for every $t \geq 1$. Also, by the taxation principle, for sufficiently large t , if the mechanism

operates on v_i, v_{-i} then it will allocate the bidder i a set $S \in \mathcal{S}$ such that $S \cap U \in B$ (provided that such a set exists).

We now split the proof into two proofs (one for Theorem 5.13 and the other for Theorem 5.14).

Theorem 5.14: We note that given, v_{-i}, \mathcal{S} and a *value query access* to the (characteristic function of the) collection B , it is possible to simulate a value query to v_i by a single value query to B . Therefore, the mechanism can be used to solve the following problem: Given a value query access to B , determine whether B is empty. It is clear that the query complexity of this last problem is $2^{|U|} - 1$ and, therefore, the query complexity of M is at least $2^{|U|} - 1 \geq 2^{m^\epsilon} - 1$.

Theorem 5.13: We note that if membership in B can be evaluated by a polynomial sized circuit, then, by Lemma 6.7, v_i can be also realized by a polynomial sized circuit. Using this fact, it is not hard to show a (nonuniform) polynomial-time algorithm for, say, MAXCUT. Given a graph on the vertices U and a number r , we will define B as the collection of all subsets $S \subset U$ such that the cut $(S, U \setminus S)$ consists of $\geq k$ edges. By running the mechanism with v_i corresponding to this B , we will be able to efficiently identify whether the graph has a cut of size $\geq k$. \square

5.2.3. Computational lower bounds on capped-additive valuations. Our next result considers the class of valuations known as capped-additive (or budget-additive), where a bidder has an additive valuation over the items, together with a restriction on the maximal total price she is willing to pay.

DEFINITION 5.15. *A valuation function v is said to be capped-additive if there exist an additive valuation a and a real value B , such that for every bundle $S \subset [m]$, $v(S) = \min\{a(S), B\}$.*

THEOREM 5.16. *Let $\epsilon > 0$. There is no polynomial-time VCG mechanism with an approximation ratio of $m^{0.5-\epsilon}$ w.r.t. capped-additive valuations unless $NP \subset P/poly$.*

This theorem has been proved in [22, 5, 7, 18], which first applied the notion of VC dimension in order to lower bound the performance of mechanisms. The proof presented here somewhat simplifies the step of finding a shattered set.

Proof. First, observe that capped-additive valuations strictly contain additive valuations. Hence, the existence of such a mechanism M together with Theorem 5.10 implies that its bank of allocations \mathcal{H} shatters a pair $A \in \binom{[n]}{2}$, $S \subset [m]$ with $|S| \geq m^{\epsilon/4}$. This allows us to embed SUBSET-SUM (see [7]), and computational hardness follows. \square

6. Inapproximability for general mechanisms. We next describe our approach to proving inapproximability results for unrestricted deterministic mechanisms. We illustrate these ideas by outlining the proof of a new separation gap between truthful and nontruthful deterministic mechanisms for combinatorial auctions with general valuations. Specifically, we show that for any choice of constant $\epsilon > 0$, there exists a class of succinctly described valuation functions such that (1) a nontruthful computationally efficient algorithm achieves an approximation ratio of m^ϵ for this class, but (2) no computationally efficient and truthful algorithm (mechanism) can obtain an $O(m^{1-\epsilon})$ -approximation unless $NP \subseteq P/poly$. While the proof of the first part is rather straightforward and relies on an explicit construction of a mechanism that achieves the desired approximation, the proof of the second part applies a combination of the direct hardness approach, introduced by Dobzinski [11] together with the VC dimension arguments. Namely, we show that a mechanism with such approximation

guarantees must offer to one of its bidders "many options" of bundles that he may potentially obtain (depending on his valuation). Applying the standard Sauer–Shelah lemma implies the existence set of items that is shattered by these options. We then describe a reduction from vertex cover to the problem of finding the optimal allocation to the bidder, as required by a truthful mechanism.

We begin by defining the specific valuation class that we will consider.

DEFINITION 6.1. A valuation $v : 2^{[m]} \rightarrow \mathbb{R}$ is called k -local if there is $T \in \binom{[m]}{k}$ such that, for all S ,

$$v(S \cap T) \geq v(S) - \frac{1}{2m^2} v([m]) .$$

For $k = k(m)$ we denote that class of k -local valuations by k -local. The corresponding bidding language will be circuits. We are now ready to state the main result of this section:

THEOREM 6.2. For any $\epsilon > 0$,

- unless $NP \subseteq P/poly$, no efficient truthful mechanism has an approximation ratio of $m^{1-\epsilon}$ w.r.t. m^ϵ -local valuations;
- there is an efficient nontruthful mechanism with an approximation ratio of $2k$ w.r.t. k -local valuations. Moreover, the auctioneer communicates with the bidders only through value queries.

We start by describing the efficient nontruthful mechanism.

Algorithm 1.

- 1: **Input:** Oracle access to k -local valuations v_1, \dots, v_n .
 - 2: **Output:** An allocation $\{S_i\}_{i \in [n]}$
 - 3: Initialize $U = [m]$ and $B = [n]$.
 - 4: **while** $U \neq \emptyset$ and $B \neq \emptyset$ **do**
 - 5: Choose $i \in B$ that maximizes $v_i(U)$.
 - 6: Allocate i a subset $S_i \subset U$ of size k with $v_i(S_i) \geq v_i(U) - \frac{1}{2m} v_i([m])$.
 - 7: Set $B := B \setminus \{i\}$ and $U := U \setminus S_i$.
 - 8: **end while**
-

LEMMA 6.3. Algorithm 1 achieves an approximation ratio of $2k$ w.r.t. k -local valuations.

Proof. We first explain how to implement step 6 in the algorithm using only value queries. Given a set U consisting of at least k elements, we need to find a set $S \in \binom{U}{k}$ with $v(S) \geq v(U) - \frac{1}{2m} v([m])$. This can be done as follows. Simply, as long as $|U| \geq k$, we throw an element $x \in U$ that maximizes $v(U \setminus \{x\})$. Since at each step we lose at most $\frac{1}{2m^2} v([m])$, the total loss is $\leq \frac{1}{2m} v([m])$.

Next, we show that the algorithm indeed achieves an approximation ratio of $2k$. Let S_1, \dots, S_n be an optimal allocation. Assume, without loss of generality that $v_1(S_1) \geq \dots \geq v_n(S_n)$ and denote $OPT = v_1(S_1) + \dots + v_n(S_n)$. Note that in the beginning of the j th iteration $[m] \setminus U$ consists of $(j - 1)k$ elements. Therefore, one of the sets $S_1, \dots, S_{(j-1)k+1}$ is contained in U . Therefore the value of the allocation at this iteration is

$$\geq v_{(j-1)k+1}(S_{(j-1)k+1}) - \frac{1}{2m} OPT \geq \frac{1}{k} \sum_{i=1}^k v_{(j-1)k+i}(S_{(j-1)k+1}) - \frac{1}{2m} OPT$$

(here, we used the fact that $v_1(S_1) \geq \dots \geq v_n(S_n)$ and that $v_i([m]) \leq OPT$ for all bidders i). Summing over all iterations, and using the fact that there are at most $\frac{m}{k}$ iterations, we conclude that the welfare of the allocation returned by the algorithm is $\geq \frac{1}{k}OPT - \frac{m}{k} \frac{1}{2m}OPT = \frac{1}{2k}OPT$. \square

Next, we prove the second part of Theorem 6.2. We begin by defining a class of valuations:

DEFINITION 6.4. A valuation $v : 2^{[m]} \rightarrow \mathbb{R}$ is called almost single minded if there is $T \subset [m]$ such that $\forall S \subset [m]$, $v(S) = 1[T \subset S] + \frac{1}{m^3}|S|$.

These valuations will serve us in order to show the existence of a big menu for one of the bidders: Given the valuations v_{-i} of all bidders except bidder i , the menu $R_{v_{-i}}$ of bidder i , is defined to be all possible bundles that may be allocated to i , that is

$$R_{v_{-i}} := \{T \subset [m] \mid \exists v_i \text{ for which } T \text{ is allocated to } i \text{ under the valuations } v_i, v_{-i}\}.$$

We recall here the *taxation principle*. Given valuations v_{-i} , there is a nondecreasing function $p_{v_{-i}} : R_{v_{-i}} \rightarrow \mathbb{R}_+$ such that the i th bidder is allocated a set that maximizes $v_i(S) - p_{v_{-i}}(S)$ over all the sets in $R_{v_{-i}}$ and pays $p_{v_{-i}}(S)$.

DEFINITION 6.5 (structured menu [15]). Given valuations v_{-i} , a number $0 \leq k \leq m$, and $0 \leq p$ we define the structured submenu $\mathcal{S}(v_{-i}, k, p)$ as all the sets $S \subset [m]$ with

- $S \in R_{v_{-i}}$,
- $|S| = k$,
- $p - \frac{1}{m^5} < p_{v_{-i}}(S) \leq p$,
- For all $T \in R_{v_{-i}}$ such strictly contains S , $p_{v_{-i}}(T) \geq p_{v_{-i}}(S) + \frac{1}{m^3}$.

LEMMA 6.6. Let M be a mechanism with an approximation ratio of $m^{1-\epsilon}$ w.r.t. almost single-minded, m^ϵ -local valuations and assume that $n \geq m^{1-\frac{1}{2}\epsilon}$. There exists a bidder i , almost single-minded m^ϵ -local valuations v_{-i} , and numbers $1 \leq k \leq m$ and $0 \leq p \leq 2$ such that $|\mathcal{S}(v_{-i}, k, p)| \geq \frac{2^{m(1/2)\epsilon/2}}{2m^6}$.

Proof. For simplicity, we assume that $n = m^{1-\frac{1}{2}\epsilon}$. Given almost single-minded valuations v_{-i} we denote by $\mathcal{S}_{v_{-i}}$ all the sets $T \subset [m]$ that the bidder i can obtain by bidding with an almost single-minded valuation.

CLAIM 3 (see [11]).

- For every $S \in \mathcal{S}_{v_{-i}}$ and every $T \in R_{v_{-i}}$ that strictly contains S , $p_{v_{-i}}(T) \geq p_{v_{-i}}(S) + \frac{1}{m^3}$.
- For every $S \in \mathcal{S}_{v_{-i}}$, $0 \leq p_{v_{-i}}(S) \leq 2$.

By this claim, it is enough to show that for some v_{-i} , $|\mathcal{S}_{v_{-i}}| \geq \frac{2^{2n}}{m}$. Indeed, in that case, by the claim, we have

$$\mathcal{S}_{v_{-i}} \subset \bigcup_{0 \leq k \leq m, 0 \leq p' \leq 2m^5} \mathcal{S}\left(v_{-i}, k, \frac{p'}{2m^5}\right).$$

By the pigeonhole principle, there is some $\mathcal{S}(v_{-i}, k, \frac{p'}{2m^5})$ consisting of $\geq \frac{2^{2n}}{2m^6}$ sets.

Suppose, toward a contradiction, that it is not the case. Consider random almost single-minded valuations v_1, \dots, v_n sampled uniformly from all such collections for which each v_i corresponds to a set S_i of size $\frac{m}{2n}$ and the sets S_i are mutually disjoint.

Let $\{T_i\}_{i \in [n]}$ be the allocation given by the mechanism operating on these random valuations. We say that the bidder i is satisfied if $S_i \subset T_i$.

CLAIM 4. *With a positive probability, none of the bidders are satisfied by a set of size $\leq \frac{m}{4}$.*

This entails contradiction since in that case there are at most 4 satisfied bidders, and therefore the total welfare is ≤ 5 . On the other hand, there is some allocation that satisfies all bidders, and therefore its welfare is $\geq m^{1-\frac{1}{2}\epsilon}$. This contradicts the assumption that the approximation ratio is $m^{1-\epsilon}$. It is therefore left to prove the claim.

Proof. It is enough to show that the probability that the n th bidder is satisfied by a set of size $\leq \frac{m}{4}$ is $< \frac{1}{n}$. Indeed, given the sets S_1, \dots, S_{n-1} , the set S_n is a random subset of $X' = X \setminus (\cup_{i=1}^{n-1} S_i)$. Since $|X'| \geq \frac{1}{2}m$, the probability that the n th bidder is satisfied by a fixed set $U \subset X$ of size $\leq \frac{n}{4}$ is at most $(\frac{1}{2})^{\frac{m}{2n}}$. Now, since $|S_{v_{-n}}| < \frac{2\frac{m}{n}}{m}$, there are less than $\frac{2\frac{m}{n}}{m}$ potential such sets, and the claim follows. \square

We proceed by applying the Sauer–Shelah lemma which provides us with a subset of items $X \subseteq [m]$ of size $|X| = \Omega(m^{\frac{1}{4}})$ that is shattered by some $\mathcal{S}(v_{-i}, k, p)$. The fact that X is shattered by a structured submenu with the same level of prices allows us to solve a hard computational problem by considering a valuation that relies on these prices. In order to define such valuation, we extend the prices from $R_{v_{-i}}$ to all subsets of $[m]$. We do so as follows: Given $T \subset [m]$, set $p_{v_{-i}}(T) := \min_{T \subset S, S \in R_{v_{-i}}} p_{v_{-i}}(S)$. We will need the following two facts.

LEMMA 6.7 (see [15]). *Let M be an efficient truthful mechanism and let v_{-i} be some valuations with polynomial description. Then membership in $\mathcal{S}(v_{-i}, k, p)$ and calculation of (the extended) $p_{v_{-i}}$ can be realized by a circuit of polynomial size.*

We are now ready to prove the second part of Theorem 6.2.

LEMMA 6.8. *An efficient truthful mechanism cannot provide an approximation ratio of $m^{1-\epsilon}$ w.r.t. m^ϵ -local valuations, unless $NP \subseteq P/poly$.*

Proof. Let M be a mechanism as specified. Fix almost single mined, m^ϵ -local valuations v_{-i} , k and $0 \leq p \leq 2$ for which $|\mathcal{S}(v_{-i}, k, p)| \geq \frac{2^{m^{(1/2)\epsilon}/2}}{2m^6}$. Denote $\mathcal{S} = \mathcal{S}(v_{-i}, k, p)$. By the Sauer–Shelah lemma, there is a set $U \subset [m]$ of size $m^{\frac{\epsilon}{4}}$ (for sufficiently large m) that is shattered by \mathcal{S} .

We will reach a contradiction by showing that this entails an efficient (non-uniform) algorithm that detect whether a graph on the vertex set U has a vertex cover of size t .

Consider the valuation $v_i = v_i^1 + v_i^2$ defined as follows: $v_i^1(S) = 6m^2$ if $S \cap U$ is a vertex cover and 0 otherwise. While

$$v_i^2(S) = \begin{cases} 3, & (S \in \mathcal{S} \text{ and } |S \cap U| = t) \text{ or } (p_{v_{-i}}(S) > p \text{ and } |S| > k), \\ 0, & \text{otherwise.} \end{cases}$$

It is not hard to see that v_i is nondecreasing, m^ϵ -local, and by Lemma 6.7 can be realized by a polynomial sized circuit. Also, by the taxation principle, it is not hard to check that if there is a cover C of size t of U , then the mechanism, operating on v_1, \dots, v_n , will allocate the bidder i a set T such that $T \cap U$ is a cover of size t . This concludes the proof. \square

7. Open questions. Inapproximability results for deterministic truthful mechanisms. As Table 1 shows, important questions about the approximability of deterministic truthful mechanisms remain open, including the approximability of

general truthful mechanisms for general and submodular valuations in the communication complexity model, and the approximability of general truthful mechanisms for capped-additive (and other classes of succinctly described) valuations in the computational complexity model.

Inapproximability results for randomized truthful mechanisms. We applied VC machinery to establish lower bounds for deterministic mechanisms. What about randomized mechanisms? Consider, e.g., the subclass of truthful mechanisms that are maximal in *distributional* range (MIDR [12]), i.e., always output the best probability distribution over outcomes in a predetermined set of probability distributions (that need not contain all distributions). Can VC machinery be extended to prove lower bounds for MIDR (as with MIR)?

Optimization over partial domains. SAT_L is an example of an interesting, and largely unexplored, class of problems in complexity theory: optimization over “partial domains,” i.e., when the space of possible solutions is (potentially) incomplete. Intuitively, any NP-hard problem remains hard so long as the space of solutions is “large.” Our results here can, from this perspective, be regarded as providing the means for proving this in the context of set-packing problems. Partial-domain variants of other classical problems seem to require new machinery. Consider, e.g., k -SET-COVER, in which the input is a collection of subsets S_1, \dots, S_n and the goal is to determine whether U can be covered by k of these subsets. What about k -SET-COVER $_L$, in which the indices of the k subsets must now belong to some large predetermined set $L \subseteq [n]^k$?

Tight bounds for the new dimension. We observe that when $k = n$ the bound of Theorem 1.5 is tight for every choice of d, m . To see this, fix $m \geq d \geq 0$ and $n \geq k \geq 2$ and consider the class $\mathcal{H} \subset [n]^{[m]}$ consisting of all functions $f : [m] \rightarrow [n]$ for which $|\{x \mid f(x) \geq k\}| \leq d$. Now, observe that $\text{Dim}_k(\mathcal{H}) = d$ and

$$|\mathcal{H}| = \sum_{i=0}^d \binom{m}{i} (k-1)^{n-i} (n+1-k)^i.$$

However, if $k < n$ the bounds derived from Theorem 1.5 are loose. Sharpening the upper/lower bounds is left for future research.

Appendix A. Hardness of the k -packing promise problem. Recall that the k -packing promise problem is the following. The input consists of r sets $S_1, \dots, S_r \subset U$ and an integer $C > 0$, and the goal is to distinguish between the following two options:

- *Positive:* There are C sets in S_1, \dots, S_n that are pairwise disjoint.
- *Negative:* For every collection \mathcal{F} of C sets out of S_1, \dots, S_n , there is some $x \in U$ that is covered by k sets of \mathcal{F} .

THEOREM A.1. *For every constant k , the k -packing promise problem is UGC-hard.*

Proof. We shall use a reduction from the following promise problem: Given a k -uniform hypergraph $G = (V, E)$, one should distinguish between the following cases:

- *Yes case:* There exists $V' \subset V$, satisfying $|V'| \geq \frac{1}{2k}|V|$, such that no hyperedge contains two or more vertices from V' .
- *No case:* Every set V' of size $\geq \frac{1}{2k}|V|$ contains an edge.

This problem (actually, a more general one) was shown to be UGC-hard in [2]. The reduction to the k -packing promise problem is given by referring to the hyperedges

as the elements of the universe. For each vertex $i \in V$ define S_i to be the set of all hyperedges that contain i . Finally, let $C = \frac{1}{2k}|V|$.

It is straightforward to verify that this defines a valid reduction to the k -packing promise problem. \square

Acknowledgments. We thank Subhash Khot for referring us to [2] and pointing out that the problem studied there can be reduced to the k -packing promise problem (see the proof of Theorem A.1). We thank Guy Kindler and Nati Linial for valuable discussions.

REFERENCES

- [1] N. ALON, S. BEN-DAVID, N. CESA-BIANCHI, AND D. HAUSSLER, *Scale-sensitive dimensions, uniform convergence, and learnability*, J. ACM, 44 (1997), pp. 615–631.
- [2] N. BANSAL AND S. KHOT, *Inapproximability of hypergraph vertex cover and applications to scheduling problems*, in Automata, Languages and Programming, Springer, Berlin, 2010, pp. 250–261.
- [3] Y. BARTAL, R. GONEN, AND N. NISAN, *Incentive compatible multi unit combinatorial auctions*, in Proceedings of the 9th Conference on Theoretical Aspects of Rationality and Knowledge, ACM, New York, 2003, pp. 72–87.
- [4] S. BEN-DAVID, N. CESA-BIANCHI, D. HAUSSLER, AND P. LONG, *Characterizations of learnability for classes of $\{0, \dots, n\}$ -valued functions*, J. Comput. System Sci., 50 (1995), pp. 74–86.
- [5] D. BUCHFUHRER, S. DUGHMI, H. FU, R. KLEINBERG, E. MOSSEL, C. PAPADIMITRIOU, M. SCHAPIRA, Y. SINGER, AND C. UMANS, *Inapproximability for VCG-based combinatorial auctions*, in Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2010, pp. 518–536.
- [6] D. BUCHFUHRER, M. SCHAPIRA, AND Y. SINGER, *Computation and incentives in combinatorial public projects*, in Proceedings of the 11th ACM Conference on Electronic Commerce, ACM, New York, 2010, pp. 33–42.
- [7] D. BUCHFUHRER AND C. UMANS, *Limits on the social welfare of maximal-in-range auction mechanisms*, in Electronic Colloquium on Computational Complexity (ECCC), Technical report, TR09-068, 2009.
- [8] W. CHEN, L. PENG, J. WANG, D. XIE, F. LI, AND M. TANG, *The complexity and algorithm for k -duplicates combinatorial auctions with submodular and subadditive bidders*, in Proceedings of the International Conference on Foundations of Computer Science (FOCS), The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), CSREA Press, Athens, GA, 2013, p. 1.
- [9] A. DANIELY AND S. SHALEV-SHWARTZ, *Optimal learners to multiclass problems*, in COLT, 2014, Proc. Mach. Learn. Res. 35, Proceedings of Machine Learning Research, 2014, pp. 287–316.
- [10] A. DANIELY, S. SABATO, S. BEN-DAVID, AND S. SHALEV-SHWARTZ, *Multiclass learnability and the ERM principle*, JMLR Workshop Conf. Proc. 19 (2011), pp. 207–232.
- [11] S. DOBZINSKI, *An impossibility result for truthful combinatorial auctions with submodular valuations*, in Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, ACM, New York, 2011, pp. 139–148.
- [12] S. DOBZINSKI AND S. DUGHMI, *On the power of randomization in algorithmic mechanism design*, in 50th Annual IEEE Symposium on Foundations of Computer Science, 2009. FOCS'09, IEEE, IEEE Computer Society, Los Alamitos, CA, 2009, pp. 505–514.
- [13] S. DOBZINSKI AND N. NISAN, *Limitations of VCG-based mechanisms*, in Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing, ACM, New York, 2007, pp. 338–344.
- [14] S. DOBZINSKI AND M. SCHAPIRA, *Optimal Upper and Lower Approximation Bounds for k -Duplicates Combinatorial Auctions*, Technical report, Hebrew University of Jerusalem, Jerusalem, 2005.
- [15] S. DOBZINSKI AND J. VONDRÁK, *The computational complexity of truthfulness in combinatorial auctions*, in Proceedings of the 13th ACM Conference on Electronic Commerce, ACM, New York, 2012, pp. 405–422.
- [16] S. DOBZINSKI, N. NISAN, AND M. SCHAPIRA, *Approximation algorithms for combinatorial auctions with complement-free bidders*, in Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing, ACM, New York, 2005, pp. 610–618.

- [17] S. DOBZINSKI, N. NISAN, AND M. SCHAPIRA, *Truthful randomized mechanisms for combinatorial auctions*, J. Comput. System Sci., 78 (2012), pp. 15–25.
- [18] S. DUGHMI, H. FU, AND R. KLEINBERG, *Amplified Hardness of Approximation for VCG-Based Mechanisms*, preprint, arXiv:0907.1948, 2009.
- [19] D. HAUSSLER AND P. M. LONG, *A generalization of Sauer’s Lemma*, J. Combin. Theory Ser. A, 71 (1995), pp. 219–240.
- [20] R. HOLZMAN, N. KFIR-DAHAV, D. MONDERER, AND M. TENNENHOLTZ, *Bundling equilibrium in combinatorial auctions*, Games Econom. Behav., 47 (2004), pp. 104–123, [https://doi.org/10.1016/S0899-8256\(03\)00184-2](https://doi.org/10.1016/S0899-8256(03)00184-2).
- [21] D. LEHMANN, L. I. O’CALLAGHAN, AND Y. SHOHAM, *Truth revelation in approximately efficient combinatorial auctions*, J. ACM, 49 (2002), pp. 577–602.
- [22] E. MOSSEL, C. PAPADIMITRIOU, M. SCHAPIRA, AND Y. SINGER, *VC v. VCG: Inapproximability of Combinatorial Auctions Via Generalizations of the VC Dimension*, preprint, arXiv:0905.1995, 2009.
- [23] B. K. NATARAJAN, *On learning sets and functions*, Mach. Learn., 4 (1989), pp. 67–97.
- [24] N. NISAN, *The communication complexity of approximate set packing and covering*, in Automata, Languages and Programming, Springer, Berlin, 2002, pp. 868–875.
- [25] N. NISAN, *Algorithmic Game Theory*, Cambridge University Press, Cambridge, 2007.
- [26] N. NISAN AND I. SEGAL, *The communication requirements of efficient allocations and supporting prices*, J. Econom. Theory, 129 (2006), pp. 192–224.
- [27] C. PAPADIMITRIOU, M. SCHAPIRA, AND Y. SINGER, *On the hardness of being truthful*, in Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS’08 2008, IEEE, Piscataway, NJ, IEEE, 2008, pp. 250–259.
- [28] N. SAUER, *On the density of families of sets*, J. Combin. Theory Ser. A, 13 (1972), pp. 145–147.
- [29] S. SHELAH, *A combinatorial problem; stability and order for models and theories in infinitary languages*, Pacific J. Math, 41 (1972), pp. 247–261.
- [30] J. M. STEELE, *Existence of submatrices with all possible columns*, J. Combin. Theory Ser. A, 24 (1978), pp. 84–88.
- [31] V. N. VAPNIK, *The Nature of Statistical Learning Theory*, Springer, New York, 1995.