Working Paper on Revenue Maximization in Generalized Deferred-Acceptance Auctions

Dimitris Fotakis, Ermis Soumalias, Panagiotis Patsilinakos
School of Electrical and Computer Engineering
National Technical University of Athens, 15780 Athens, Greece
Author1, Author2,Author3@corelab.ntua.gr

***Abstract*—We study the problem of creating a weakly group-strategyproof (WGSP) mechanism for the very natural setting of a Multi-Unit auction. We study both the case where bidders have linear valuations, and the harder case of decreasing marginal values. In the first case, we give an upper bound on the sample complexity of achieving a $(1-\epsilon)$ approximation to the optimal expected revenue. In the case of decreasing marginal values, no WGSP mechanism can guarantee an approximation ratio better than $\sqrt{2}$ even for the much simpler problem of maximizing Social Welfare. With the class of WGSP mechanisms that we propose however, under mild distribution assumptions, given logarithmic in the number of players and units available for sale samples, one can determine a mechanism with expected revenue close to the expected optimal social welfare. We then extend both our results for the cases where there are respectively only 1 or 2 samples available.**

I. Introduction

Multi-unit auctions appear in many different contexts (e.g., spectrum auctions [22], treasury auctions [21], IPO auctions [20] etc.) and have been studied extensively (and virtually from every possible aspect) for a few decades, starting from Vickrey's seminal paper (1961). The general case is NP-hard, as the single-minded case is just a reformulation of the knapsack problem. For this reason, as well as their multitude of applications, special cases of multi-unit auctions are usually studied, making assumptions such as knowledge of the bidders' valuation distribution or properties of the bidders' valuations (most commonly "downward-sloping valuations"). A different point of view, towards which a lot of recent results are oriented, is sample-based mechanism design ([1],[3], [17], [18], [19]). In these settings, the mechanism designer has no knowledge of the distribution itself, but instead he has access to some independent samples from it. The main motivation of this is the fact that full knowledge of the distribution is often highly impractical. In recent works ([1],[2],[3], [17]) generalization guarantees were provided, bounding the difference between the empirical profit of a mechanism over a set of samples and its expected profit on the unknown distribution.

Another rapidly developing area of study emerged from the work of Milgrom and Segal [11], which proposed the framework of Deferred-Acceptance (DA) Auctions, a family of auctions with a number of remarkable incentive properties. Gkatzelis, Markakis and Roughgarden [5] generalized the DA Auction framework to non-binary settings, maintaining its attractive properties: Truthful bidding is a dominant strategy, even in the sense of *obvious strategyproofness* formalized by Li [13]. Additionally, every DA auction satisfies weak group-strategyproofness: No coalition of bidders can submit false bids in a way that strictly increases the utility of every bidder of the coalition. We refer the reader to [14] for further advantages and motivation behind DA Auctions. Unfortunately, DA Auctions do not always achieve good approximation to social welfare. Multi-unit auctions are being deployed in practice in a wide variety of applications, and more recently they have also been used by various online brokers (eg [15],[16]). In the multi-unit setting Gkatzelis, Markakis and Roughgarden [5] proved that when bidders have decreasing marginal values no WGSP mechanism can guarantee an approximation ratio for social welfare better than $\sqrt{2}$. For more results on the power and limitations of DA auctions, see the work of Dutting et al. [23].

In this work, we define a different family of mechanisms for the multi-unit auction setting: This family consists of all the mechanisms that choose prior to the actual auction the bundle of units that will offer to each bidder, and the corresponding price. In the actual auction, every bidder is only offered a bundle of that predetermined size, at the predetermined posted price. If his bid is above his posted price, he is allocated that bundle of units, at that reserve price. Otherwise, he is allocated no units. We denote that mechanism class $\mathcal{UB}$. Mechanisms of this class can trivially be implemented as DA auctions, satisfying the requirements introduced in [5], thus they possess the same attractive characteristics. We show that, under mild assumptions about the bidders' valuations, using a logarithmic to the number of players and units available number of samples, one can determine a mechanism M from the proposed class, $\mathcal{UB}$, whose expected revenue is close to the expected optimal social welfare. Specifically, we firstly analyze the case of linear bidders, utilizing a modification of the *t-level auctions* introduced by *Morgenstern and Roughgarden* [1]. Then, we study the much harder case of submodular bidders, using the framework provided in the recent work of Balcan et al. ([2]). Finally, we propose a mechanism, aptly named *Two Samples Mechanism*, that using only 2 samples, achieves expected revenue close to $1/2$ of the expected revenue of the optimal mechanism of the $\mathcal{UB}$ class. This final analysis is motivated by the fact that weakly group-strategyproof mechanisms are often used in high-impact markets, selling expensive, large scale goods such as radio spectrum (e.g. [11],[12],[14]). Such auctions are quite often rare so assuming even a logarithmic, non constant number of samples might be unrealistic.

A. The setting

We consider mechanisms that are ex post incentive compatible and individually rational, meaning that it is a dominant strategy for agents to participate in the mechanism and bid their true private values, weakly group strategy-proof (WGSP), meaning no coalition of bidders can collectively submit false bids in a way that makes every bidder of the coalition strictly better off, and obviously strategyproof, meaning that from a behavioural perspective, even cognitively limited agents can recognize truthful bidding as a weakly dominant strategy. A basic example would be selling identical goods with unlimited supply to agents at a fixed price per unit. Every bidder chooses how many units he wants, and pays the respective price.

We are interested in the setting of a m-unit auction with bidder-specific demands, where each bidder wants up to a specific number of units, and his value increases with the number of units he is allocated. The set of feasible allocations is precisely the ones that allocate at most m units. We focus on the typical case of submodular bidders, i.e. the extra value that a bidder gains for an additional unit weakly decreases with the number of units that that bidder has already clinched.

As is standard in the Bayesian optimal auction literature, we assume that each bidder's marginal values are drawn independently from unknown distributions. A standard and common assumption on such distribution is that they are regular, a condition equivalent to concavity of the revenue as a function of the probability of sale (for a single item and a single bidder). An important subclass of regular distributions are those meeting the monotone hazard rate (MHR) condition: intuitively, these have tails no heavier than that of an exponential distribution (which has constant hazard rate). Uniform, normal, and exponential distributions meet the MHR condition.

B. The main ideas

All our sample complexity results are proved in two parts. The interface between the two is the a priori optimal mechanism for social welfare, i.e. the mechanism that with perfect distribution knowledge chooses the fixed allocation with the highest expected social welfare, and performs that allocation with no consideration of the bidders' actual values.

In the first part of our proof, we establish a lower bound for the expected revenue of the optimal mechanism from the proposed class, compared to the expected social welfare of the a priori optimal mechanism. In the case of submodular bidders, this is accomplished using Balcan's Sample Comlexity Framework [2]. In the case of linear bidders, this is accomplished using a slight modification of Hartline and Rougharden's Simple Versus Optimal Mechanisms [7].

The second part of our proof is the same in both cases. We prove a lower bound on the expected social welfare of the a priori optimal mechanism compared to the expected social welfare of the Vickrey-Clarke-Groves (VCG) mechanism. This is the only part where we use our distribution assumptions, implying that the similar results may hold for different distribution assumptions.

II. Preliminaries

A *mechanism* comprises an allocation and a payment rule. An *allocation rule* $\mathbf{x}$ is a function from bid profiles to $\mathbb{N}^n$, indicating how many units each bidder is allocated. A *payment rule* $\mathbf{p}$ is a function from bid profiles to $\mathbb{R}^n$, indicating the payment of each bidder. A mechanism is *truthful* if for every bidder i and every fixed set v_{-i} of bids by the other bidders, the bidder i maximizes its *utility* $\mathrm{u}_i(\mathrm{v}) = v_i(x_i(\mathrm{v})) - p_i(\mathrm{v})$ by bidding its true valuation (as opposed to some other false bid). This paper studies only truthful mechanisms, and thus we do not distinguish between bidders' true valuations and the bids they submit to the mechanism. A mechanism is weakly group-strategyproof (WGSP) if for any coalition of bidders $S \subseteq N$, for any vector $\boldsymbol{b}_{-S} = (b_j)_{j \notin S}$ and for any vector $\boldsymbol{b'}_S = (b'_j)_{j \in S}$ of the bidders in S, it holds that

$$\mathrm{u}_i(\mathrm{v}_S, \boldsymbol{b}_{-S}) \geq \mathrm{u}_i(\boldsymbol{b'}_S, \boldsymbol{b}_{-S}) \text{ for some } i \in S.$$

Hence, there is no coalition that can make all its members strictly better off by deviating from the truth. All our mechanisms are obviously strategy proof. A strategy is *obviously dominant* if, for any deviation, at any information set where both strategies first diverge, the best outcome under the deviation is no better than the worst outcome under the dominant strategy. A mechanism is *obviously strategy-proof* (OSP) if it has an equilibrium in obviously dominant strategies. Finally a mechanism is *individually rational* if $p_i(\mathrm{v}) \leq v_i(x_i(\mathrm{v}))$ for every bidder i and input v, implying that truthful bidders are guaranteed non-negative utility by the mechanism.

The *Vickrey-Clarke-Groves* (VCG) mechanism in our setting works as follows: Given the valuations of the agents, the mechanism selects an allocation x to maximize the *social welfare*, i.e. $\sum_{i=1}^{n} v_i(x_i(\mathrm{v}))$ subject to feasibility. For every bidder i, his total payment for the units he is allocated is the lowest total amount he could have bid (for that amount of units) at which it would continue to win that exact amount of units, which is the difference between the surplus of other agents in the optimal allocation and the surplus of an optimal allocation that excludes bidder i. One can easily check that the VCG mechanism is truthful and individually rational.

A bidders' *demand* is the maximum number of units that interest him. Allocating to a bidder more units that his demand does not increase that bidders value. For every bidder, his *marginal value* is his increase in value for gaining an additional unit, given the units he already has. We call a bidder *linear* if his marginal value remains constant up to his demand, i.e. his value for a number of units is his value per unit times the number of units he is allocated (up to his demand). A bidder is called *submodular* if his marginal value is weakly decreasing with the number of units he is allocated. In the case of linear bidders, their valuations are drawn from a product distribution $F = F_1 \times \cdots \times F_n$, and in the case of submodular bidders, from a product distribution $F = F_{1,1} \times \dots F_{1,m} \times F_{2,1} \times \cdots \times F_{n,m}$. As such, the bidders marginal values are independently, but not identically, distributed. For simplicity, we assume that every distribution F_i has a continuous density function f_i that is strictly positive on the support of the distribution, which we assume is an interval of the non-negative real line. We denote the support of F using the notation $\mathcal{X}$. The *hazard rate* of the distribution F_i at a point z in its support is defined as $h_i(z) = f_i(z)/(1 - F_i(z))$. A distribution F satisfies the *monotone hazard rate* (MHR) condition if its hazard rate is nondecreasing over its support. Many of the most common distributions (exponential, uniform, normal, etc.) satisfy this condition. A weaker condition is *regularity*, which requires only that $z - 1/h(z)$ is nondecreasing over the support of the distribution and thereby allows for heavier tails. A canonical distribution that is regular but not MHR is the *equal-revenue* distribution, defined by $F(z) = 1 - 1/z$ on $[1, \infty)$.

Our objective is to maximize the revenue of the mechanism. The revenue of a mechanism $\mathcal{M} = (\boldsymbol{x}, p)$ on an input v, denoted by Rev($\mathcal{M}$, v), is the sum of the payments collected: $\sum_{i=1}^{n} p_i(\mathrm{v})$. Generally, two different mechanisms earn incomparable revenue: one will collect more on some inputs, the other on other inputs. However, for a fixed distribution over valuations the expected revenues of different mechanisms are absolutely comparable. As is traditional in optimal mechanism design, we assume that the mechanism designer has some information on the bidders' distributions, but not on their actual values.

We say that a mechanism class $\mathcal{M}$ is (d, t)-delineable if:

1) The class $\mathcal{M}$ consists of mechanisms parametrized by vectors $\boldsymbol{p}$ from a set $\mathcal{P} \subseteq \mathbb{R}^d$ and
2) For any $\mathrm{v} \in \mathcal{X}$, there is a set $\mathcal{H}$ of t hyperplanes such that any connected component $\mathcal{P}'$ of $\mathcal{P} \setminus \mathcal{H}$, the function $\mathrm{Rev}(\mathcal{M}_p, \mathrm{v})$ is linear over $\mathcal{P}'$.

Let $S = \{\mathrm{v}^{(1)}, \mathrm{v}^{(2)}, \dots, \mathrm{v}^{(N)}\}$ be a subset of $\mathcal{X}$ and let $\mathrm{z}^{(1)}, \mathrm{z}^{(2)}, \dots, \mathrm{z}^{(N)} \in \mathbb{R}$ be a set of *targets*. We say that $\mathcal{X}$ and let $\mathrm{z}^{(1)}, \mathrm{z}^{(2)}, \dots, \mathrm{z}^{(N)}$ witness the shattering of S by $\mathcal{M}$ if for all $T \subseteq S$, there exists some mechanism $M_T \in \mathcal{M}$ such that for all $\mathrm{v}^{(i)} \in T$, $Rev(M_T, \mathrm{v}^{(i)}) \leq z^{(i)}$ and for all $\mathrm{v}^{(i)} \notin T$, $Rev(M_T, \mathrm{v}^{(i)}) > z^{(i)}$. If there is some $\boldsymbol{z} \in \mathbb{R}^N$ that witnesses the shattering of S by $\mathcal{M}$, then we say that S is *shatterable* by $\mathcal{M}$. Finally, the pseudo-dimension of $\mathcal{M}$, denoted Pdim($\mathcal{M}$), is the size of the largest set that is shatterable by $\mathcal{M}$.

Theorem 2.1 (Pollard 1984): For any mechanism class $\mathcal{M}$, let U be the maximum profit achievable by mechanisms in $\mathcal{M}$. Then:

$$\epsilon_{\mathcal{M}(N,\delta)} = O\left(U\sqrt{\mathrm{Pdim}(\mathcal{M})/N} + U\sqrt{\ln(1/\delta)/N}\right),$$

where $\epsilon_{\mathcal{M}}(N, \delta)$ is a function such that the following generalization guarantee holds:
With probability $1 - \delta$ over the draw $S \sim F^N$, for any mechanism $M \in \mathcal{M}$,

$$|\mathrm{Rev}(M, S) - \mathrm{Rev}(M, F)| \leq \epsilon_{\mathcal{M}}(N, \delta),$$

where $\mathrm{Rev}(M, S)$ is the empirical revenue of M on S, i.e. its average revenue on the sample, and $\mathrm{Rev}(M, D)$ is the expected revenue of M on distribution D.

Theorem 2.2 (Balcan et al. 2018): If $\mathcal{M}$ is (d,t)-delineable, the pseudo-dimension of $\mathcal{M}$ is $O(d \log(dt))$.

III. Linear Bidders - Sample Complexity Bounds

In this section, we provide an upper bound on the sample complexity of choosing a Mechanism whose expected revenue is a $(1-\epsilon)$-approximation of the expected revenue of Myerson, for all cases where the bidders' valuations are linear, i.e. for every bidder i his value for acquiring s_i units of the good is $l_i \cdot v_i$ where v_i is his value per unit. There are 2 distinct cases: when the bidders' demands are bounded, and when they are not. The reason for this distinction is that in the case of unbounded demands, we can prove the same complexity bound using weaker distribution assumptions. A technical highlight of this section would be that we extend the class of t-level matroid auctions to the class of t-level multi-unit auctions, and show that the same sample complexity bounds still hold. This may be of independent interest.

A. Unbounded Demand

We examine the following setting:

- n bidders
- multi-unit auction with m units available for sale
- each bidder is *additive*: For every bidder i, his valuation for acquiring l_i units of the good is $l_i \cdot u_i$, where u_i: his value per unit of the good.
- each bidder i's value per unit, v_i, follows a distribution with probability density function $f_i(\cdot)$, with virtual valuation function $\phi_i(\cdot)$.
- The distributions of all bidders' values per unit of the good are either bounded on [1,H] or Monotone Hazard Rate(MHR). In this setting only, we can show that the same bounds hold for MHR distributions, instead of bounded ones.

The mechanism is very similar to the n-bidder, single item t-level auction introduced in [1]. The main differences are that now we use that mechanism on the value per unit, and that we convert the allocation and payment rules to that of a Deferred-Acceptance auction, which immediately implies weak group-strategyproofness.

Each bidder i faces t thresholds: $0 \leq l_{i,0} \leq l_{i,1} \leq \cdots \leq l_{i,t-1}$. This set of $t \cdot n$ numbers defines a Linear DA t-level auction as described in LDA-C_t.

Theorem 3.1: For a fixed tie-breaking ordering $\prec$, the pseudo-dimension of the set of n-bidder linear DA t-level Auctions is $O(nt \log(nt))$.

Proof. The only observation necessary is that the only winner in this auction is the highest-level bidder, and in the case of ties, the one favoured by the lexicographical ordering $\prec$. If we focus on a single unit, this is the exact same allocation rule as the single-item t-level auctions introduced in [1]. Since

Algorithm 1 Linear DA t-level Auction (LDA-C_t)

1: For each bidder i, let $t_i(v_i)$ denote the index of the largest threshold $l_{i,\tau}$ that lower bounds v_i (or -1, if $v_i < l_{i,0}$). We call $t_i(v_i)$ the level of bidder i.

2: Sort the bidders from lowest level to highest and, within a level, use a fixed lexicographical tie-breaking ordering $\prec$ to pick a winner.

3: DA Allocation: Start finalizing the bidders, from lowest level to highest. The first n-1 bidders to be finalized aren't allocated any units. The last bidder to be finalized is allocated all m units.

4: Payment rule: The payment rule is the unique one that renders truthful bidding a dominant strategy and charges 0 to losing bidders (same as the payments in the t-level Auctions for the single item case in [1], multiplied by m for the m units).

both that mechanism and the one we propose use Myersonian payments, the 2 auctions also have the same payment rule. The two auctions are equivalent and thus their proof [1, Theorem 3.3] holds as is.

Lemma 3.2: Consider n bidders with valuations (per unit) in [0,H] and with $\mathbb{P}[max_i v_i > \alpha] \geq \gamma$. Then, LDA$-C_t$ contains a DA Auction with expected revenue at least $1-\epsilon$ times that of an optimal auction, for $t = \left(\frac{1}{\gamma\epsilon} + \log_{1+\epsilon}\frac{H}{\epsilon}\right)$.

Proof. The proof is very similar to that of [1, Lemma 3.5]. For the sake of completeness, we include it here, along with all necessary changes.

Consider a fixed bidder i. We define t thresholds for i, bucketing i by his virtual value, and prove that the t-level DA Auction $\mathcal{A}$ induced by these thresholds for each bidder closely approximates the expected revenue of the optimal auction $\mathcal{M}$. Let ϵ' be a parameter to be defined later.

1) Set $l_{i,0} = \phi^{-1}(0)$, bidder i's monopoly reserve.
2) For $\tau \in [1, \lceil 1/\gamma\epsilon' \rceil]$, let $l_{i,\tau} = \phi_i^{-1}(\tau \cdot \alpha\gamma\epsilon)$ ($\phi_i \in [0,1]$)
3) For $\tau \in [\lceil 1/\gamma\epsilon' \rceil, \lceil 1/\gamma\epsilon' \rceil + \lceil \log_{1+\frac{\epsilon}{2}} \frac{H}{a} \rceil]$, let $l_{i,\tau} = \phi_i^{-1}(\alpha \cdot (1+\frac{\epsilon}{2})^{\tau - \lceil \frac{1}{\gamma\epsilon'} \rceil})$ $(\phi_i > 1)$

Consider a fixed valulation profile v. Let i^* denote the winner according to $\mathcal{A}$ and let $i^{'}$ denote the winner according to $\mathcal{M}$. We note that $\mathcal{M}$ always awards all the units to bidders with the highest positive virtual value per unit, or to no one, if no such bidder exists. The definition of the thresholds immediately implies the following:

1) $\mathcal{A}$ only allocates to non-negative ironed virtual valued-bidders.
2) If there is no tie (i.e.there is a unique bidder at the highest level), then $i^* = i^{'}$.
3) If there is a tie at level τ, the virtual value per unit of the winner of $\mathcal{A}$ is close to that of $\mathcal{M}$:
   - If $\tau \in [0, \lceil 1/\gamma\epsilon' \rceil]$, then $\phi_{i'}(v_{i'}) - \phi_{i^*}(v_{i^*}) \leq \alpha\gamma\epsilon'$.
   - If $\tau \in [\lceil 1/\gamma\epsilon' \rceil, \lceil 1/\gamma\epsilon' \rceil + \lceil \log_{1+\frac{\epsilon}{2}} \frac{H}{a} \rceil]$, then $\frac{\phi_{i^*}(v_{i^*})}{\phi_{i'}(v_{i'})} \geq 1 - \frac{\epsilon}{2}$.

These facts imply that:

$$\begin{aligned}\mathbb{E}_{\mathrm{v}}[Rev(\mathcal{A})] &= \mathbb{E}_{\mathrm{v}}[m \cdot \phi_{i^*}(v_{i^*})] = m \cdot \mathbb{E}_{\mathrm{v}}[\phi_{i^*}(v_{i^*})] \\ &\geq m\left[\left(1-\frac{\epsilon}{2}\right)\mathbb{E}_{\mathrm{v}}[\phi_{i'}(v_{i'})] - \alpha\gamma\epsilon'\right] \\ &\geq \left(1-\frac{\epsilon}{2}\right) \cdot \mathbb{E}_{\mathrm{v}}[m \cdot \phi_{i'}(v_{i'})] - m\alpha\gamma\epsilon' \\ &= \left(1-\frac{\epsilon}{2}\right)\mathbb{E}_{\mathrm{v}}[Rev(\mathcal{M})] - m\alpha\gamma\epsilon'\end{aligned}$$

Where the second and third inequalities follow from linearity of expectations. The assumption that $\mathbb{P}[max_i v_i > \alpha] \geq \gamma$ implies (since a feasible solution sets price α for each unit and awards all units to any bidder with valuation per unit at least α) that $\mathbb{E}_{\mathrm{v}}[Rev(\mathcal{M})] \geq m\alpha\gamma$. Combining this with (1), and setting $\epsilon' = \frac{\epsilon}{2}$ implies $\mathbb{E}_{\mathrm{v}}[Rev(\mathcal{A})] \geq \left(1-\frac{\epsilon}{2}\right) \cdot \mathbb{E}_{\mathrm{v}}[Rev(\mathcal{M})]$. □

By taking $\alpha = \gamma = 1$ we get:

Corollary 3.2.1: Suppose F is a product distribution over $[1,H]^n$. If $t = \Omega\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$ then LDA-C_t contains a DA auction with expected revenue at least $1-\epsilon$ times the expected optimal revenue.

Corollary 2.2.1 and [1, Theorem 2.1] immediately yield the following corollary:

Corollary 3.2.2: Let F be a product distribution with all bidders' valuations in [1,H]. Assume that $t = \Theta\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$ and $N = O\left(\left(\frac{H}{\epsilon}\right)^2\left(nt\log(nt)\log\frac{H}{\epsilon} + \log\frac{1}{\delta}\right)\right) = \widetilde{O}\left(\frac{H^2 n}{\epsilon^3}\right)$. Then with probability at least $1-\delta$, the empirical revenue maximizer of LDA-C_t on a set of N samples from F has expected revenue at least $(1-\epsilon)$ times that of the optimal auction.

In a very similar vein, we can slightly edit the proof of Section 4 of [1] to generalize Corollary 2.2.1 for MHR distributions. Thus:

Corollary 3.2.3: With probability $1-\delta$, the empirical revenue maximizer of a sample of size N of the class of t-level η-truncated Linear DA Auctions is a $(1-O(\epsilon))$-approximation of the optimal auction for n MHR bidders, for $t = O\left(\frac{1}{\epsilon'} + \log_{1+\epsilon'}\left(\log\frac{1}{\epsilon'}\right)\right)$, $\epsilon' = \frac{\epsilon}{\log\frac{1}{\epsilon}}$ and $N = O\left(\left(\frac{1}{\epsilon'}\right)^2\left(nt\log(nt)\ln\frac{1}{\epsilon'} + \ln\frac{1}{\delta}\right)\right) = \widetilde{O}\left(\frac{n}{\epsilon^3}\right)$, where η can be learned fromm the set of the N samples

B. Bounded Demand

We examine the following setting:

- n bidders
- multi-unit auction with m units available for sale
- every bidder i has a publicly known demand d_i
- every bidder is additive, up to his demand d_i:
  For every bidder i, his value for acquiring l_i units of the good (a service level l_i), up to d_i, is $l_i \cdot v_i$, where v_i: His value per unit of the good. For more than d_i units: his value remains $v_i \cdot d_i$.

- Every bidder i's value per unit, v_i, follows some distribution with probability density function $f_i(\cdot)$ and virtual valuation function $\phi_i(\cdot)$
- All n bidders have values per unit in $[1, H]$.
(their distributions / valuations are bounded)

For our arguments to hold, the demands of the players need to be known in advance of the auction, so as the environment remains single-parameter.

The mechanism class is similar to the t-level matroid auctions introduced in [1]. Similarly with the first case, we converted the mechanism to a multi-unit DA auction to ensure weak group-strategyproofness. Compared to the unbounded demand case, there is one key difference: Now the proposed mechanism is not equivalent to a t-level auction of the ones introduced in [1]. Instead, we extend their matroid t-level auction to a multi-unit t-level auction, and prove that the same sample complexity and approximation bounds still hold.

Algorithm 2 Bounded Demand, Linear DA t-level Auctions (BLDA)

1: For each bidder i, let $t_i(v_i)$ denote the index of the largest threshold $l_{i,\tau}$ that lower bounds v_i (or -1, if $v_i < l_{i,0}$). We call $t_i(v_i)$ the level of bidder i.
2: Sort the bidders from lowest level to highest and, within a level, use a fixed lexicographical tie-breaking ordering $\prec$ to pick a winner.
3: DA Allocation: Start finalizing the bidders, from lowest level to highest.
(equivalently: $\sigma_i^{A_t}(b_i, b_{N\setminus A_t}) = t_i(b_i)$, and in the case of ties, finalize the bidders in the same as the fixed lexicographical ordering$\prec$)
4: Clinching function:
$g_i^{A_t}(b_{N\setminus A_t}) = \min\left\{d_i, m - \min\left\{m, \sum_{j\in A_t\setminus\{i\}} d_j\right\}\right\}$
(Every bidder finalized is allocated, up to his demand, as many units as possible, provided the units left afterwards are enough to fully satisfy the demand of the still active players)
5: They payment rule is the unique one that renders the mechanism truthful and charges 0 to bidders who win no items (Myerson payments).

Lemma 3.3: This is a valid DA Auction.

Proof. The scoring function of every bidder depends only on his bid (so it doesn't depend on the bids of the other still active players) and is weakly increasing, since $\forall i \quad 0 \leq l_{i,0} \leq l_{i,1} \leq \dots \leq l_{i,t-1}$. The clinching function is non-increasing with respect to the set of active players, i.e. $g_i^{A_{t+1}}(b_{N\setminus A_{t+1}}) \geq g_i^{A_t}(b_{N\setminus A_t})$. According to [5] in any generalized single-parameter environment, any auction that satisfies those constraints is a valid DA Auction. □

Theorem 3.4: The pseudo-dimension of the class of Bounded Demand, Linear DA t-level auctions with n bidders is $O(nt\log(nt)$.

The proof is similar to [1, Theorem 5.1], but some modifications are required. We'll need 2 standard results from learning theory:

Lemma 3.5 (Sauer's Lemma): Let $\mathcal{C}$ be a set of functions from $\mathcal{Q}$ to {0,1} with VC dimension d, and $\mathcal{S} \subseteq \mathcal{Q}$. Then:

$$|\{\mathcal{S} \cap \{x \in \mathcal{Q} : c(x) = 1\} : c \in \mathcal{C}\}| \leq |S|^d$$

Lemma 3.6: The set of linear separators in $\mathbb{R}^d$ has VC dimension $d + 1$.

Now we can proceed with the proof of Theorem 2.4:

Proof. Consider a set of samples S of size N which can be shattered by bounded demand, linear DA t-level auctions with revenue targets $(r^1, r^2, \dots, r^N)$. We upper-bound the number of labelings of S possible using bounded demand, linear DA t-level auctions, which yields an upper bound on N.

We partition the auctions into equivalence classes, identically to the proof of [1, Theorem 3.3]. Across all auctions in an equivalence class, all comparisons between two thresholds or a threshold and a bid are resolved identically. From [1], the number of equivalence classes are at most $(nm + nt)^{2nt}$. We now upper bound the number of distinct labelings any fixed equivalence class $\mathcal{C}$ of auctions can generate.

Consider a class $\mathcal{C}$ of equivalent auctions. The allocation and payment rules are relatively simple. The number of units a bidder i wins depends only on 3 things: the ordering of the bidders (by level), the fixed tie-breaking rule $\prec$ and the demands of all bidders and thus, for fixed demands, is a function only of bids and thresholds. This implies that, for every sample in S, all auctions in the class $\mathcal{C}$ result in the exact same allocation (number of units allocated to each bidder). This, along with Myerson payments, also imply that the payment of each winning bidder is a fixed sum of thresholds and coefficients that correspond to how many additional items he clinched because of that threshold:

$$p_{\mathcal{C},i}(\mathrm{v}) = \sum_{j \leq t(v_i)} l_{i,j} \cdot h_{\mathcal{C},i}(v, j),$$

where: $h_{\mathcal{C},i}(v, j)$: How many additional units bidder i clinches in all auctions in $\mathcal{C}$, for the valuation profile v, because he is at least at level j. For any bidder i, the function $h_{\mathcal{C},i}()$ is the same across all samples in set S, for any auction in $\mathcal{C}$.

Now, we encode each $\mathcal{A} \in \mathcal{C}$ and sample v^j as an $(nt+1)$-dimensional vector as follows. Let $x_{i,\tau}^{\mathcal{A}}$ encode the value of $l_{i,\tau}$ in the auction $\mathcal{A}$. Define $x_{nt+1}^{\mathcal{A}} = 1$ for every $\mathcal{A} \in \mathcal{C}$. Define $y_{i,\tau}^j = h_{\mathcal{C},i}(v, j)$ (for how many units bidder i is paying his τ-th threshold).Finally, define $y_{nt+1} = -r^j$. The point is that, for every auction $\mathcal{A}$ in the class $\mathcal{C}$ and sample v^j,

$$x^{\mathcal{A}} \cdot y^j \geq 0$$

if and only if $\mathrm{Rev}(\mathcal{A}, \mathrm{v}^j) \geq r^j$. Thus, the number of distinct labelings of the samples generated by auctions in $\mathcal{C}$ is bounded above by the number of distinct sign patterns on N points in $\mathbb{R}^{nt+1}$ generated by all linear separators (The y^j-vectors are constant across $\mathcal{C}$ and can be viewed as m fixed points in $\mathbb{R}^{nt+1}$; each auction $\mathcal{A} \in \mathcal{C}$ corresponds to the vector $x^{\mathcal{A}}$ of coefficients.). Applying Sauer's Lemma and Lemma 3.6, BLDA t-level auctions can generate at most N^{nt+2} labelings per equivalence class, and hence at most $(nN + nt)^{3nt+2}$ distinct labelings in total. This imposes the restriction

$$2^n \leq (nN + nt)^{3nt+2}$$

solving for N yields the desired bound.

□

Remark: The proof is very similar to the proof of [1, Lemma 5.3]. The key difference between the 2 proofs is that now the vector y^j that corresponds to the payment identity of the auctions in $\mathcal{C}$ on the j-th sample doesn't consist only of {0,1} in its first nt positions (which threshold each winner must pay), but instead those positions are in $\mathbb{N}$. Now, the position $y^j_{i,k}$ corresponds to how many times the bidder must pay his k-th threshold (because he could have clinched more than 1 units because of that threshold). However, the rest of the argument can remain exactly the same: Again, for any sample v^j the payment identity vector y^j is the same for all auctions in the equivalence class $\mathcal{C}$, and can be viewed as a fixed point in $\mathbb{R}^{nt+1}$. The auctions can be encoded in the exact same way as in the original proof, and so we can use the same arguement as in the original proof, where each auction is equivalent to the coefficients of a linear separator.

Remark 2: It is interesting that the demands of the bidders don't have to be the same across all samples (if you notice the proof, it is not used anywhere), just publicly known in advance for every sample. Essentially they have to not be part of the players' bids so that we remain on a single-parameter environment.

A more intuitive explanation is that the demand vector, for every sample, only constrains the possible allocations on that sample, so the possible payment identities y^j; every y^j must satisfy the polymatroid constraint that for every i, its coordinates $y_{i,0}$ to $y_{i,t-1}$ sum up to at most d_i. But the proof only uses that those payment identities y are fixed for any equivalence class $\mathcal{C}$. It makes no use of the added constraint imposed by the demands.

Now, all we have to show is a low representation bound, i.e. that there is at least one mechanism in the proposed class with high expected revenue.

Theorem 3.7: Consider the environment described above (linear bidders, bounded demands). Suppose F is a product distribution with support $[1, H]^n$. Provided $t = \Omega\left(\frac{1}{\epsilon} + \log_{1+\epsilon} H\right)$, there exists a BLDA t-level auction with expected revenue at least a $1 - \epsilon$ fraction of the optimal expected revenue.

Proof. The proof is similar to [1, Theorem 5.4], but again, some non-trivial modifications are required.

Define the bidders' thresholds exactly as in our own Lemma 3.2 and let $\mathcal{A}$ denote the corresponding BLDA t-level auction. Fix an arbitrary valuation profile v. We'll compare the Virtual Social Welfare of $\mathcal{A}$ to that of $\mathcal{M}$, where $\mathcal{M}$ is the Myerson Auction. For every unit allocated to some bidder i, up to his demand d_i, the contribution of that unit to the virtual social welfare is $\phi_i(v_i)$ (i.e. equal to that bidder's virtual value).

By its definition, $\mathcal{M}$ allocates units that correspond to the m highest duplicated virtual values, where each virtual value $\phi_i(v_i)$ has been duplicated d_i times (the demand of the corresponding bidder), provided those m duplicated virtual values are positive.

In a similar vein, the allocation of $\mathcal{A}$ is lexicographically optimal with respect to the levels, rather than the exact virtual values. Since the allocation of $\mathcal{M}$ is the actual optimal, it is also optimal with respect to the t-levels. Thus, the level of the unit with the i-th highest virtual value in the allocation of $\mathcal{A}$ is the same as in of $\mathcal{M}$ (otherwise $\mathcal{A}$ wouldn't be optimal with respect to the levels).

The 2 mechanisms always allocate the same number of units: the minimum between m and the number of duplicated virtual values ≥ 0. Then, by an accounting argument identical to [1,Theorem 3.5] summing up over all units completes the proof:

- Consider a fixed valuation profile v.
- Let i^* and i' denote the i-th highest unit in the allocation of $\mathcal{A}$ and $\mathcal{M}$ respectively.
- Both auctions only allocate to non-negative ironed virtual values.
- Let τ be the level of i^*
- If there is no tie at that level, then both Mechanisms allocate that same unit.
- When there is a tie at level τ, the virtual value of the unit allocated by $\mathcal{A}$ is close to that of $\mathcal{M}$:
  - If $\tau \in [0, \lceil 1/\gamma\epsilon' \rceil]$, then $\phi_{i'}(v_{i'}) - \phi_{i^*}(v_{i^*}) \leq \alpha\gamma\epsilon'$.
  - If $\tau \in [\lceil 1/\gamma\epsilon' \rceil, \lceil 1/\gamma\epsilon' \rceil + \lceil \log_{1+\frac{\epsilon}{2}} \frac{H}{a} \rceil]$, then $\frac{\phi_{i^*}(v_{i^*})}{\phi_{i'}(v_{i'})} \geq 1 - \frac{\epsilon}{2}$.

These facts imply that:

$$\begin{aligned}
\mathbb{E}_{\text{v}}[\text{Rev}_i(\mathcal{A})] &= \mathbb{E}_{\text{v}}[\phi_{i^*}(v_{i^*})] \\
&\geq \left[\left(1 - \frac{\epsilon}{2}\right) \mathbb{E}_{\text{v}}[\phi_{i'}(v_{i'})] - \alpha\gamma\epsilon'\right] \\
&\geq \left(1 - \frac{\epsilon}{2}\right) \cdot \mathbb{E}_{\text{v}}[\phi_{i'}(v_{i'})] - \alpha\gamma\epsilon' \\
&= \left(1 - \frac{\epsilon}{2}\right) \mathbb{E}_{\text{v}}[\text{Rev}_i(\mathcal{M})] - \alpha\gamma\epsilon'
\end{aligned}$$

Summing over the first m units, for both of the auctions, completes the proof.

□

IV. Submodular Bidders - Unit Bundling

Now we tackle the problem of creating a WGSP multi-unit auction for submodular bidders. This setting has some additional difficulties. Namely, the environment is no longer single-parameter, thus we can not compare the expected revenue of our mechanism against Myerson. We'll need to compare the expected revenue of the proposed mechanism directly against the expected social welfare of VCG. This introduces our second difficulty: No WGSP mechanism can guarantee an approximation ratio better than $\sqrt{2}$ even for the simpler problem of Social Welfare maximization.

A. Bounding the Pseudodimension of The Unit Bundling Mechanism Class

This subsection introduces the proposed mechanism class and bounds its pseudo-dimension using Balcan's Sample Complexity framework.

Algorithm 3 Unit Bundling Mechanism Class ($\mathcal{UB}$)

1: Have n sizes $s_1, s_2, \ldots, s_n$ and n prices $p_1, p_2, \ldots, p_n$.
2: Offer each bidder i the bundle containing s_i units at price r_i per unit, so $r_i \cdot s_i$ in total.

Theorem 4.1: $\mathcal{UB}$ is $(2n, 3nm)$-delineable.

Proof. For the parameter space $\mathcal{P}$ we have:

- $(r_1, r_2, \ldots, r_n) \in \mathbb{R}^n_+$
- $(s_1, s_2, \ldots, s_n) \in \mathbb{N}^n$
- $s_1 + s_2 + \cdots + s_n \leq m$.

Fix a valuation profile v. Let $\mathrm{v}_i(1), \mathrm{v}_i(2), \ldots, \mathrm{v}_i(m)$ be bidder i's marginal values.We need to check $2m$ hyperplanes of the form $s_i \geq \alpha, si \leq \alpha, \alpha \in [m]$ to determine how many units are included in the bundle of bidder i (technicality, but I think it's neeeded). Once we have determined the answer $\widetilde{\alpha}$ to that question, the profit from bidder i is simply:

$$\text{profit}_{i,p}(\cdot) = \begin{cases} 0, \text{if } \mathrm{v}_i(1) + \mathrm{v}_i(2) + \ldots, \mathrm{v}_i(\widetilde{\alpha}) < \widetilde{\alpha} \cdot r_i \\ \widetilde{\alpha} \cdot r_i, \text{ else} \end{cases}$$

So, once we have determined $\widetilde{\alpha}$, there is just one hyperplane splitting the parameter space $\mathcal{P}$ in regions where the profit from bidder i is linear: $\mathrm{v}_i(1) + \mathrm{v}_i(2) + \ldots, \mathrm{v}_i(\widetilde{\alpha}) \geq \widetilde{\alpha} \cdot r_i$.

In total, we have m options for α in our parameter space ($\alpha \leq m$) and then for each such choice, there is one additional hyperplane splitting the parameter space in 2 subregions where the profit from bidder i is linear. Thus, the total number of hyperplanes splitting the parameter space in regions where the profit from bidder i is linear is $2m + m = 3m$.

Repeating the same process for all bidders, there are at most $n \cdot 3m$ hyperplanes splitting the parameter space in connected components, in each of which the profit from all bidders is linear. Thus, the total profit, which is just the sum of the profit from each bidder, will also be linear in each one of those connected components of $\mathcal{P}$.

□

Now that we've showed that the $\mathcal{UB}$ class is $(2n, 3nm)$-delineable, applying [2, Theorem 2.2] gives us an upper bound on its pseudo-dimension.

Corollary 4.1.1: The pseudo-dimension of $\mathcal{UB}$ is $O(n \log(nm))$.

B. Bounding the empirical revenue of $\mathcal{UB}$

This subsection shows that the expected revenue of the optimal mechanism from the $\mathcal{UB}$ class is close to the expected social welfare of the a priori optimal mechanism, $\mathcal{A}$. This subsection, together with A, make up the first part of our proof.

For any mechanism $\mathcal{M}$, its revenue on a valuation profile v can be viewed as:
Rev($\mathcal{M}$,v) = SW($\mathcal{M}$, v) - S($\mathcal{M}$, v), where:

- SW($\mathcal{M}$, v): The social welfare of $\mathcal{M}$ on v.
- S($\mathcal{M}$, v): The separation of the mechanism $\mathcal{M}$ on v, i.e. how much money it "left on the table" for the bidders / the social welfare that it didn't extract as revenue.

On any valuation profile v, the revenue of any mechanism $\mathcal{M}$ of the class UB is simply the sum of its revenue for every bidder:

$$\begin{aligned} \text{Rev}(\mathcal{M}, \mathrm{v}) &= \text{Rev}_i(\mathcal{M}, \mathrm{v}) = \\ &= \sum_{i=1}^{n} [\text{SW}_i(\mathcal{M}, \mathrm{v}) - \text{S}_i(\mathcal{M}, \mathrm{v})], \quad \text{where:} \end{aligned}$$

- $\text{SW}_i(\mathcal{M}, \mathrm{v})$: Bidder i's contribution to the Social Welfare, i.e. bidder i's value in the allocation of $\mathcal{M}$ for the valuation profile v.
- $\text{S}_i(\mathcal{M}, \mathrm{v})$: The separation of bidder i, i.e. his utility/ how much of bidder i's value $\mathcal{M}$ didn't extract as revenue.

We'll compare the revenue of the optimal mechanism from $\mathcal{UB}$ against the Social Welfare achieved by the a priori optimal mechanism for social welfare, $\mathcal{A}$. $\mathcal{A}$ is a hypothetical mechanism. It has perfect distribution knowledge, and chooses the fixed allocation with the highest expected social welfare, without ever considering the true bidders' valuations. It has a very simple description:

Algorithm 4 A priori optimal mechanism ($\mathcal{A}$)

1: Sort all marginal values in descending order according to their expected values. (Ties are broken lexicographically).
2: Allocate to every bidder as many items as his marginal values in the first m positions of that list.

Theorem 4.2: Let $\mathcal{M}$ be the optimal mechanism from the $\mathcal{UB}$ class. For the expected revenue of $\mathcal{M}$, it holds:

$$\text{Rev}_F(\mathcal{M}) \geq \text{SW}_F(\mathcal{A}) - \sum_{i=1}^{n} (3B_i)^{1/3} \widetilde{\sigma}_i^{2/3} \tag{1}$$

where:

- $\text{Rev}_F(\mathcal{M})$: The expected revenue that $\mathcal{M}$ extracts on the distribution F.
- B_i: Bidder i's expected value for s_i units $B_i = \sum_{j=1}^{s_i} \mathbb{E}_F[\mathrm{v}_{i,j}]$
- $\widetilde{\sigma}_i^{\,2}$: The variance of bidder i's value distribution for receiving s_i units:

$\widetilde{\sigma_i}^2 = \sum_{j=1}^{s_i} \mathrm{Var}(F_{i,j}) = \sum_{j=1}^{s_i} \sigma_{i,j}^2$

Proof. Mechanism $\mathcal{A}$ allocates exactly m units, and the amount of units allocated to each bidder i, x_i, is chosen (for a fixed product distribution of the bidders' marginal valuations, F) in advance. There is a mechanism $\mathcal{M}$ in $\mathcal{UB}$ that chooses bundle sizes $s_1 = x_1, s_2 = x_2, \ldots, s_n = x_n$ and sets the optimal prices (per unit) for those bundles $r_1, r_2, \ldots, r_n$. We'll compare the expected revenue of $\mathcal{M}$ against the expected social welfare of $\mathcal{A}$.

Once the bundle sizes are fixed, for each bidder $\mathcal{M}$ faces a distinct posted price problem. From [3, Corollary 2] we have that for any distribution D:

$$S \leq (3B)^{1/3}\sigma^{2/3}, \text{ where :} \tag{2}$$

- S : The expected separation for the optimal reserve price
- B : The mean of D.
- σ^2: The variance of D.

Applying (2) to the posted price problem the mechanism $\mathcal{M}$ faces against a fixed bidder i, we have:

$$\mathrm{Rev}_{i,F}(\mathcal{M}) = B_i - (3B_i)^{1/3} \cdot \widetilde{\sigma_i}^{2/3}$$

where $Rev_{i,F}(\mathcal{M})$ is the expected revenue of $\mathcal{M}$ from bidder i, when the bidders' valuations are drawn from distribution F.

Summing up over all bidders and using linearity of expectations, for the expected revenue of $\mathcal{M}$ we have:

$$\mathrm{Rev}_F(\mathcal{M}) = \sum_{i=1}^{n} \mathrm{Rev}_{i,F}(\mathcal{M}) \geq \sum_{i=1}^{n} [B_i - (3B)^{(1/3)} \cdot \widetilde{s_i}^{2/3}]$$

$$\implies \mathrm{Rev}_F(\mathcal{M}) \geq \mathrm{SW}_F(\mathcal{A}) - \sum_{i=1}^{n} (3B_i)^{1/3} \widetilde{\sigma_i}^{2/3}$$

□

V. Bounding the difference $\mathrm{SW}_F(\mathrm{VCG}) - \mathrm{SW}_F(\mathcal{A})$

For any valuation profile v, the allocation of $\mathcal{A}$ can be turned into the allocation of VCG using at most m unit "moves":
"move": Remove a unit from the allocation of $\mathcal{A}$ that doesn't correspond to a marginal value that the VCG used and allocate it to a marginal value that VCG did use, but $\mathcal{A}$ did not.

For any distribution F, the difference $\mathrm{SW}_F(\mathrm{VCG}) - \mathrm{SW}_F(\mathcal{A})$ is exactly the expected sum of the gain in social welfare by those (at most m) moves.

We define an ordering on those moves:
move(i): For a given valuation profile v, remove the unit that corresponds to the i-th lowest expected value, out of the marginal values that $\mathcal{A}$ did allocate an item to, but VCG did not, and allocate that unit where it corresponds to the i-th highest expected value, out of the marginal values that VCG did allocate an item to, but $\mathcal{A}$ did not.

Let gain(move(i)) denote the gain in Social Welfare from move(i). It is obvious from the definition of the moves that:

$$\mathbb{E}_F[gain(move(i))] \geq \mathbb{E}_F[gain(move(i+1))] \ \ \forall i \in [m-1],$$

and

$$\mathrm{SW(VCG)} - \mathrm{SW}(\mathcal{A}) = \sum_{i=1}^{m} \mathbb{E}_F[\text{gain(move(i))}] \tag{A}$$

A. Bounding (A) for Gaussian/Sub-Gaussian distribution

We'll make some additional distribution assumptions. These assumptions are only used to upper bound the expected gain of each move. Similar bounds can be proven for other kind of assumptions.

1) All marginal values are drawn exclusively from Gaussian and Sub-Gaussian distributions.
2) Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{mn}$ be their expected values, in sorted order.
3) Finally, we assume that all standard deviations/ Sub-Gaussian parameters can be sorted in the same relative way, $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_{mn}$ where (μ_i, σ_i) correspond to the same distribution.
(This is a weaker assumption that $\sigma_i \sim \mu_i$ or $\sigma_i^2 \sim \mu_i$, i.e. standard deviation or variance proportional to expected value.)

Theorem 5.1: Under these assumptions, for the expected gain of the i-th move it holds:

$$\mathbb{E}_F[gain(move(i))] \leq \mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} + \sigma_1\sqrt{2\log m} \tag{3}$$

Proof. The lowest possible expected value of the distribution of the unit we are removing from the allocation of $\mathcal{A}$ is μ_{m+1-i} because by the definition of move(i), for the units that correspond to lower expected marginal values in the allocation of $\mathcal{A}$, if they were to be removed, they would have been removed by previous moves.

The highest possible expected value of the distribution of the unit we are adding, from the allocation of VCG that wasn't in the allocation of $\mathcal{A}$ is μ_{m+i} because again, by the definition of move(i), if units were to be allocated to marginal values with expected value in $\{ \mu_{m+1}, \mu_{m+2}, \ldots, \mu m + i - 1\}$, they would have been allocated by previous moves.

When each move(i) is allocating a unit to some marginal value, it only has $(n-1)m$ options, because m out of the mn marginal values are already included in the allocation of $\mathcal{A}$. Because both the mean and standard deviation/Sub-Gaussian parameter decrease as we "move down" in the two sorted lists:

- If all the $(n-1)m$ marginal values not included in the allocation of $\mathcal{A}$ were sampled from Gaussian/Sub-Gaussian distributions with parameters $(\mu_{m+i}, \sigma_{m+i}^2)$ (the highest possible), their expected maximum would be at most $\mu_{m+i} + \sigma_{m+i}\sqrt{2\log((n-1)m)}$.
- If any of the $(n-1)m$ marginal values follows a distribution with either smaller mean or standard deviation, the expected value of their maximum strongly decreases.

- So, the expected value of the item that the i-th move is adding to the allocation of $\mathcal{A}$ is at most $\mu_{m+i} + \sigma_{m+i}\sqrt{2\log((n-1)m)}$.

Similarly, when each move(i) is removing a unit from the allocation of $\mathcal{A}$, it only has m options, because $\mathcal{A}$ allocates at most m units. Because both the mean and standard deviation/Sub-Gaussian parameter decrease as we "move-down" in the two sorted lists:

- The maximum possible variance of the distribution of a unit in the allocation of $\mathcal{A}$ is σ_1.
- If all the m marginal values in the allocation of $\mathcal{A}$ were sampled from Gaussian/Sub-Gaussian distributions with parameters $(\mu_{m+1-i}, \sigma_1^2)$, their expected minimum would be at least $\mu_{m+1-i} - \sigma_1\sqrt{2\log(m)}$.
- If any of the m marginal values follows a distribution with greater mean, the expected value of their minimum strongly increases.
- Thus the expected value of the item that the i-th move is removing from the allocation of $\mathcal{A}$ is at least $\mu_{m+1-i} - \sigma_1\sqrt{2\log(m)}$.

Since the expected gain of every move is simply the value of the unit it is adding minus the value of the one it's removing the two paragraphs above immediately imply:

$$\mathbb{E}_F[gain(move(i))] \leq \mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} + \sigma_1\sqrt{2\log m} \quad \text{(B)}$$

Combining (A) and (B) with the fact that on any valuation profile, any one of the m moves that will actually will be performed has non-negative gain, we have:

$$\text{SW}_F(\mathcal{A}) = \text{SW}_F(\text{VCG}) - \sum_{i=1}^{m} \max\{\mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} + \sigma_1\sqrt{2\log m}, 0\} \quad \text{(C)}$$

VI. Combining Parts 1 and 2 of our Proof

In this section we provide all the necessary arguments to combine parts 1 and 2 of our proof, showing that with probability $1-\delta$ over the draw of the sample S, the expected revenue of the empirically optimal mechanism of the $\mathcal{UB}$ class will be close to the expected optimal social welfare.

Theorem 6.1: In the setting of submodular bidders whose marginal values follow exclusively Gaussian/Sub-Gaussian distributions and the ordering of the expected values of those distributions is the same as the ordering of their standard deviations/Sub-Gaussian parameters, with probability $1-\delta$ over the draw of the sample $S \sim F^N$, for the empirically optimal mechanism $\mathcal{M}$:

$$\begin{aligned}\text{Rev}_F(\mathcal{M}) \geq\ & \text{SW}_F(\text{VCG}) \\ & - \sum_{i=1}^{m} \max\{\mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} \\ & + \sigma_1\sqrt{2\log m},\ 0\} \\ & - \sum_{i=1}^{n} (3B_i)^{1/3}\widetilde{\sigma}_i^{\,2/3} \\ & - O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right)\end{aligned}$$

Proof. So far, using the framework of [2], we upper-bounded the pseudo-dimension of UB-N as $O(n\log(nm))$. From the classic work of [4], this immediately implies that with probability $1-\delta$ over the draw of $S \sim F^N$, for any $\mathcal{M} \in$ UB-N:

$$|\text{Rev}_S(\mathcal{M}) - \text{Rev}_F(\mathcal{M})| \leq O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right) \quad (4)$$

,where:

- $\text{Rev}_S(\mathcal{M})$: The empirical revenue of $\mathcal{M}$ on the sample set S
- $\text{Rev}_F(\mathcal{M})$: The expected revenue of $\mathcal{M}$ on the distribution F
- U: The maximum profit achievable by any mechanism in the UB class.

Then, we showed using [3] that there exists a mechanism $\mathcal{M} \in \mathcal{UB}$ such that:

$$\text{Rev}_F(\mathcal{M}) \geq \text{SW}_F(\mathcal{A}) - \sum_{i=1}^{n} (3B_i)^{1/3}\widetilde{\sigma}_i^{\,2/3} \quad (5)$$

(4) and (5) immediately imply that with probability at least $1-\delta$ over the draw of the sample set $S \sim F^N$, for the empirically optimal mechanism $\mathcal{M}$ of $\mathcal{UB}$:

$$\begin{aligned}\text{Rev}_F(\mathcal{M}) \geq\ & \text{SW}_F(\mathcal{A}) - \sum_{i=1}^{n} (3B_i)^{1/3}\widetilde{\sigma}_i^{\,2/3} \quad (6) \\ & - O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right)\end{aligned}$$

Combining 6 and (C), we have that in the special case of Gaussian/Sub-Gaussian distributions, with probability at least $1-\delta$ over the draw of the sample set of size N, for the empirically optimal mechanism $\mathcal{M}$ determined on the sample:

$$\begin{aligned}\text{Rev}_F(\mathcal{M}) \geq\ & \text{SW}_F(\text{VCG}) \\ & - \sum_{i=1}^{m} \max\{\mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} \\ & + \sigma_1\sqrt{2\log m},\ 0\} \\ & - \sum_{i=1}^{n} (3B_i)^{1/3}\widetilde{\sigma}_i^{\,2/3} \\ & - O\left(U\sqrt{n\log(nm)/N} + U\sqrt{\ln(1/\delta)/N}\right)\end{aligned}$$

□

VII. Linear Bidders - Single Sample

In this section we tackle the same setting with linear bidders as in section 3, but now using only 1 sample. Now we do not need to make any distinctions between the Bounded and Unbounded demand cases. The same bound holds even if we have a mixture of bidders with bounded and unbounded demands, as long as all bidders are linear and their demands are publicly known in advance and not part of the players' bids. A technical highlight from this section would be that we extended the concept of *matroid commensurate* mechanisms, introduced in [7] to multi-unit/polymatroid environments, and then used that extension to prove the critical theorem of this section.

This more general setting can be stated formally:

- n bidders
- multi-unit auction with m units available for sale
- each bidder i's value per unit, v_i, follows a distribution with probability density function $f_i(\cdot)$ and virtual valuation function $\phi_i(\cdot)$.
- The demand of each bidder (up to how many units each bidder wants) can be anything: Unit-Demand, Bounded Demand, Unbounded Demand, as long as it is known in advance to the mechanism designer and not part of the players' bids.
- every bidder is additive, up to his demand d_i:
  For every bidder i, his value for acquiring l_i units of the good (a service level l_i), up to d_i, is $l_i \cdot v_i$, where v_i: His value per unit of the good. For more than d_i units: his value remains $v_i \cdot d_i$.

First, we'll have to define some necessary terms:

Duplication of a single-parameter environment:
Each bidder i with valuation distribution F_i is replaced by a pair i,i', whose valuations are i.i.d. draws from F_i. The feasible allocations in the duplicated environment are those satisfying:

1) At most one bidder from each pair is allocated any units.
2) The allocation, when naturally interpreted as an allocation in the original environment (ie all units allocated to either i or i' are allocated to i) is a feasible allocation in that environment.

(This was first defined in [7], we just include their definition for the sake of completeness)

multi-unit/polymatroid commensurate:
Let $\mathcal{M}$ and $\mathcal{M}'$ be two mechanisms for a given environment. Let $x_i(\mathrm{v})$ and $x'_i(\mathrm{v})$ denote the allocation of units to bidder i with the valuation profile v. The mechanism $\mathcal{M}$ is polymatroid commensurate with $\mathcal{M}'$ if:

$$\mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} x_i(\mathrm{v})\phi_i(v_i)] \geq 0 \tag{C1}$$

$$\mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} p_i(\mathrm{v})] \geq \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} x'_i(\mathrm{v})\phi_i(v_i)] \tag{C2}$$

where $p_i(v)$: The payment of bidder i on $\mathcal{M}$ with valuation profile v, for $x_i(\mathrm{v})$ units.

Lemma 7.1: If a mechanism $\mathcal{M}$ is polymatroid commensurate with a mechanism $\mathcal{M}'$, then

$$\mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] \geq \frac{1}{2} \cdot \mathbb{E}_{\mathrm{v}}[\mathcal{M}'(\mathrm{v})]$$

where $\mathcal{M}(\mathrm{v}), \mathcal{M}'(\mathrm{v})$: The revenue of the mechanisms $\mathcal{M}, \mathcal{M}'$ on the valuation profile v respectively.

Proof. We have:

$$\begin{aligned}\mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] &= \mathbb{E}_{\mathrm{v}}[\sum_{i\in N} x_i(\mathrm{v})\phi_i(v_i)] \\ &= \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})=x'_i(\mathrm{v})} x_i(\mathrm{v})\phi_i(v_i) + \sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} x_i(\mathrm{v})\phi_i(v_i)] \\ &= \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})=x'_i(\mathrm{v})} x_i(\mathrm{v})\phi_i(v_i)] + \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} x_i(\mathrm{v})\phi_i(v_i)] \\ &\geq \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})=x'_i(\mathrm{v})} x_i(\mathrm{v})\phi_i(v_i)] \\ &= \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})=x'_i(\mathrm{v})} x'_i(\mathrm{v})\phi_i(v_i)] \implies\end{aligned}$$

$$\mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] \geq \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})=x'_i(\mathrm{v})} x'_i(\mathrm{v})\phi_i(v_i)] \tag{7}$$

Where the only inequality follows from condition (C1). Also:

$$\begin{aligned}\mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] &= \mathbb{E}_{\mathrm{v}}[\sum_{i\in N} p_i(\mathrm{v})] \\ &\geq \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} p_i(\mathrm{v})] \\ &\geq \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} x'_i(\mathrm{v})\phi_i(v_i)] \implies\end{aligned}$$

$$\mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] \geq \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} x'_i(\mathrm{v})\phi_i(v_i)] \tag{8}$$

Where the first inequality follows from the non-negativity of payments, and the final inequality follows from the condition (C2).

From (7) and (8):

$$\begin{gathered}\mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] + \mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] \geq \\ \geq \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})=x'_i(\mathrm{v})} x'_i(\mathrm{v})\phi_i(v_i)] + \mathbb{E}_{\mathrm{v}}[\sum_{i:x_i(\mathrm{v})\neq x'_i(\mathrm{v})} x'_i(\mathrm{v})\phi_i(v_i)] \\ = \mathbb{E}_{\mathrm{v}}[\mathcal{M}'(\mathrm{v})] \implies \\ \mathbb{E}_{\mathrm{v}}[\mathcal{M}(\mathrm{v})] \geq \frac{1}{2}\mathbb{E}_{\mathrm{v}}[\mathcal{M}'(\mathrm{v})]\end{gathered}$$

□

Lemma 7.2: In the multi-unit auction with linear bidders and regular valuation distributions, VCG with duplicates is multi-unit commensurate with the optimal mechanism, Myerson, without duplicates.

Proof. We begin with the first requirement, (C1) of our definition. Let $x(\mathrm{v},\mathrm{v}'), x'(\mathrm{v},\mathrm{v}')$ denote the allocation of the VCG mechanism (with duplicates) and the optimal mechanism (without duplicates) respectively. By definition, $x'(\mathrm{v},\mathrm{v}')$ is independent of v', and can not allocate any items to duplicated bidders: $x'_{i'}(v,v')=0$.

Condition on v but not on v'; this fixes $x'(\mathrm{v},\mathrm{v'})$. We argue that:

$$\mathbb{E}_{\mathrm{v'}}[\sum_{i:x_i(\mathrm{v},\mathrm{v}')\neq x'_i(\mathrm{v},\mathrm{v}')} x_i(\mathrm{v})\phi_i(v_i)] \geq 0 \tag{9}$$

The unconditional probability in (C1) follows. We prove (9) by showing that the expected combined contribution of each original bidder i and his duplicate i' to the left-hand side is non-negative. If one of i,i' has positive allocation, it is the bidder with higher valuation and hence, by regularity, with higher virtual valuation.

First consider an original bidder i such that $x'_i(\mathrm{v},\mathrm{v}')>0$. Since the valuation distributions are regular, the optimal mechanism selects only bidders with a non-negative virtual valuation, so $\phi_i(v_i)\geq 0$. It follows that the contribution from i,i' to the left-hand side of (1) in this case is non negative with probability 1: if $x_i(\mathrm{v},\mathrm{v'})=x_{i'}(\mathrm{v},\mathrm{v'})=0$ the contribution to the virtual social welfare is zero; otherwise it is

$$\begin{aligned}\max\{x_{i'}(\mathrm{v},\mathrm{v}'),x_i(\mathrm{v},\mathrm{v}')\}\cdot\max\{\phi_i(v_i),\phi_i(v_{i'})\}\\ \geq \max\{x_{i'}(\mathrm{v},\mathrm{v}'),x_i(\mathrm{v},\mathrm{v}')\}\cdot\phi_i(v_i)\geq 0\end{aligned}$$

Now suppose that for the original bidder i we have $x'_i(\mathrm{v},\mathrm{v'})=0$. Condition further on the valuations v'_{-i} of all duplicates other than i', and let $\mathcal{E}$ denote the event that either $x_i(\mathrm{v},\mathrm{v}')>0$ or $x_{i'}(\mathrm{v},\mathrm{v}')>0$ (both of these can not occur at the same time). If $\neg\mathcal{E}$ occurs, then the contribution from i,i' to the left-hand side of (1) is zero. Since $\mathrm{v},\mathrm{v}'_{-i}$ are fixed, event $\mathcal{E}$ occurs if and only if $v_{i'}$ is at least some non-negative threshold t. In this case, the expected contribution of i,i' is $\mathbb{E}_{v_{i'}}[\max x_i(\mathrm{v},\mathrm{v}')\phi_i(v_i), x_{i'}(\mathrm{v},\mathrm{v}')\phi_i(v_{i'})]$, conditioned on $\mathrm{v},\mathrm{v}'_{-i}$ and $\mathcal{E}$. This is lower bounded by the analogous conditional expectation of $x_{i'}(\mathrm{v},\mathrm{v}')\phi_i(v_{i'})$, which is equivalent to:

$$E_{v_{i'}}[x_{i'}(\mathrm{v},\mathrm{v}')\phi_i(v_{i'})|\phi_i(v_{i'})\geq t] \tag{10}$$

Since the unconditional expectation of a virtual valuation is zero, ϕ_i is non-decreasing by regularity and $x_{i'}(\mathrm{v},\mathrm{v}')\geq 0$, the quantity in (10) is non-negative. Taking expectations over whether or not $\mathcal{E}$ occurs, and then over v'_{-i} completes the argument.

For the proof of condition (C2), we'll have to improvise. We will prove the condition pointwise, for each valuation profile v,v'.

For any valuation profile, VCG in the duplicated environment allocates at least as many units as the optimal Mechanism in the original environment, since the allocation of the original environment is always possible in the duplicated environment, and VCG chooses a maximal solution with respect to social welfare (and no negative values exist).

Focus on the bidder i with the highest virtual value per unit $\phi_i(v_i)$ to whom the optimal mechanism in the original environment allocated more units than VCG with reserves did in the duplicated environment, i.e. $x'_i(\mathrm{v},\mathrm{v}')>x_i(\mathrm{v},\mathrm{v}')$.

Since $x'_i(\mathrm{v},\mathrm{v}')>x_i(\mathrm{v},\mathrm{v}')$, allocating an additional unit to bidder i was also possible on the duplicated environment (the demands of all pairs of original/duplicated bidders are the same). So, for any unit that was allocated differently by the VCG with reserves, the externality that allocation caused was at least v_i.

By the definition of VCG, for every unit allocated differently, for the payment p the winner of that unit has to make we have: $p\geq v_i\geq\phi_i(v_i)$, where $v_i\geq\phi_i(v_i)$ follows by the definition of a virtual valuation.

Summing up over all units allocated differently by the VCG with duplicates (and using the fact that it always allocates at least as many units as the optimal mechanism in the original environment), we have that the total payments procured for those units by the VCG with reserves is greater or equal the sum of virtual values collected by the optimal mechanism, for all units that it allocated differently than the VCG with duplicates. This, combined with the fact that for the units that were allocated in the same way both mechanisms earn payments equal to the corresponding (same) virtual value completes the proof.

□

Now we can prove the main theorem of this section:

Theorem 7.3: In any setting of the above category (where all bidders have linear valuations), when bidders' valuations per unit are drawn independently from regular distributions, there exists a generalized DA Auction that, using only a single sample, has expected revenue that is a 4-approximation to the expected revenue of the optimal Auction.

Proof. The proof is similar to [7, Theorem 4.4]. Unfortunately, we can not apply that theorem directly because we are not restricted to a simple matroid setting. The Optimal Mechanism in our setting, with respect to revenue has a simple algorithmic description:

Lemmata 7.1 and 7.2 combined imply that in the multi-unit auction setting with linear bidders and regular valuation distributions, VCG with duplicates has expected revenue at least 1/2 of the expected revenue of the optimal mechanism in the original environment.

If we have a sample from the bidders' distributions, s, we can use that sample as the duplication of our environment, and run VCG on the duplicated environment v,s

where v: The valuation profile in the actual auction.

In this duplicated environment, symmetry dictates that the "winner" (bidder who is allocated any units) for any pair i,i' of bidders is equally likely to be an original bidder or a

Algorithm 5 Multi-unit Myerson for Linear Bidders

1: Collect all bids, where each bid b_i is the value of player i for a single unit.
2: For all players, calculate their virtual values per unit of the good, $\phi_i(b_i)$.
3: Sort the players according to their virtual value per unit, in descending order.
4: Start descending that list: Every player gets the minimum number of units between his demand and the number of units still available to the mechanism.
5: Charge each player the unique Myerson payments that render truthful bidding a dominant strategy and charge 0 to players that win no units.

duplicate. The expected revenue from the original bidders is thus at least a quarter of that of the optimal auction in the original environment.

□

Lemma 7.4: In this setting, the VCG mechanism is implementable as a generalized DA Auction.

Proof. We can restrict our attention to mechanisms than never allocate, to any bidder, more units than his demand: There is no loss in social welfare from doing so, because every bidder has the same valuation for receiving either as many units as his demand, or more.

The set of feasible outcomes is then defined via the submodular function $h : 2^n \to \mathbb{R}_+$, $h(S) = \min\left\{m, \sum_{i\in S} d_i\right\}$ as follows:

$$P_h = \left\{ l\epsilon\mathbb{N}^n | \sum_{i\epsilon S} l_i \leq h(S)\ \forall S \subseteq N \right\}$$

(Do not allocate to any subset of bidders more than m units, and do not allocate to any subset of bidders more units than their total demand).

It has already been proven in [1, Proposition 4.1] that when the set of feasible allocations is defined by a polymatroid constraint, the VCG mechanism is a generalized DA auction. So, in our setting, the VCG is also implementable as a generalized DA Auction. In fact, from [1, Proposition 4.1], the clinching and scoring functions of this mechanisms are as follows:

- The polymatroid auction scoring function is $\sigma_i^{A_t}(b_i, b_{N\setminus A_t}) = b_i$.
- The polymatroid auction clinching function is $g_i^{A_t}(b_{N\setminus A_t}) = h(A_t) - h(A_t \setminus \{i\})$.

Theorem 7.3 and Lemma 7.4 immediately yield the following corollary:

Corollary 7.4.1: In any setting of the above category, there exists a generalized DA Auction that, using a single sample, achieves expected revenue at least a $\frac{1}{4}$-fraction of the optimal expected revenue.

VIII. Submodular Bidders - Few Samples

In this final section we create a WGSP mechanism that, using only 2 samples, achieves expected revenue close to a $\frac{1}{2}$ of the optimal social welfare. There are 2 distinct cases: When all marginal values follow MHR distributions, and when they do not.

A. MHR Distributions

We examine the following setting:

- n bidders
- multi-unit auction with m units available for sale
- each bidder i is submodular: His "marginal value" for clinching additional units weakly decreases with the number of units he has clinched, i.e.: $v_{i,1} \geq v_{i,2} \geq \cdots \geq v_{i,m}$, where $v_{i,j}$: bidder i's value for clinching his j-th unit, provided he has already clinched j-1 units.
- All marginal values, $v_{i,j}$, follow a distribution with probability density function $f_{i,j}(\cdot)$.
- All those distributions satisfy the monotone hazard rate (MHR) condition.

Remark: If a bidder is only interested in $j < m$ units, his last $m - j$ marginal values will all be 0, with probability 1. This distribution trivially satisfies the MHR condition.

The mechanism we suggest has a simple algorithmic description:

Algorithm 6 The Two Samples Mechanism (TSM)

1: Collect 2 samples of all bidders' marginal values.
2: Run VCG on the first sample. Let $x = (x_1, x_2, \ldots, x_n)$ be the allocation of that VCG.
3: In the auction, offer to each bidder i a bundle of size x_i, at a price equal to that bidder i's value for x_i units in the second sample.

Essentially, we use one sample to determine the size of the bundle offered to each bidder, and another to determine the price of each bundle.

Lemma 8.1: When the bidders' marginal valuation distributions are MHR, the expected revenue of the two samples mechanism is at least a $\frac{1}{4}$-fraction of the expected Social Welfare of the allocation x, where x is defined as in the Two Samples Mechanism.

Proof. Let F_{x_i} denote the total distribution of bidder i's value $(v_i(x_i))$ for receiving x_i units, with x_i the bundle size determined by the Two Samples Mechanism. F_{x_i} satisfies the Monotone Hazard Rate condition, because the random variable $v_i(x_i)$ is the sum of x_i random variables, all of which are MHR.
(Very short proof: MHR distributions have tails that decay at least as fast as exponential distributions. The sum of subexponential distributions is subexponential).

Against bidder i, the Two Samples Mechanism faces a posted price problem. Let $\widehat{R}_i(p) = p(1 - F_{x_i}(p))$ be the revenue Function for player i, i.e. the expected revenue the Two Samples Mechanism will earn from player i, if it offers him a bundle of size x_i at price p. Applying [8, Lemma 3.13] for $t = 0$, we have:

$$\mathbb{E}_{v_i(x_i)}[\widehat{R}_i(\max\{0, v_i(x_i)\})] \geq \frac{1}{4} \cdot V(0) \implies$$
$$\mathbb{E}_{v_i(x_i)}[\widehat{R}_i(v_i(x_i))] \geq \frac{1}{4} \cdot \mathbb{E}[v_i(x_i)]$$

where $\mathbb{E}_{v_i(x_i)}[\widehat{R}_i(\max\{0, v_i(x_i)\})] = \mathbb{E}_{v_i(x_i)}[\widehat{R}_i(v_i(x_i))]$ follows from the non-negativity of bidders' values and $V_i(0)$, the expected Social Welfare of allocating the bundle to bidder i only when his value for it is greater or equal to zero, is equal to his expected value for the bundle, because we are effectively always allocating it.

So far we have shown that for any bidder i, the expected revenue of the Two Samples Mechanism from that bidder is at least a $\frac{1}{4}$-fraction of that bidder's expected value for x_i items. Summing up over all bidders and using linearity of expectations, we have that the expected revenue of the mechanism is at least a $\frac{1}{4}$-fraction of the expected Social Welfare of the allocation x, defined as in the Two Samples mechanism.

□

Up to this point, we have shown that the Expected Revenue of the Two Sample Mechanism is at least a $\frac{1}{4}$-fraction of the Expected Social Welfare of the allocation x. Now, we need to show how close the Expected Social Welfare of x is to the Expected Social Welfare of the VCG mechanism. For this comparison, we'll make some additional assumptions about the distribution. These assumptions are the same that we used in our sample complexity bounds. The reason we are making these assumptions separately is to leave open the possibility of extending our results for different sets of distribution assumptions.

This set of additional assumptions is:

1) All marginal values are drawn exclusively from Gaussian and Sub-Gaussian distributions.
2) Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{mn}$ be their expected values, in sorted order.
3) Finally, we assume that all standard deviations/ Sub-Gaussian parameters can be sorted in the same relative way, $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_{mn}$ where (μ_i, σ_i) correspond to the same distribution.
   (This is a weaker assumption that $\sigma_i \sim \mu_i$ or $\sigma_i^2 \sim \mu_i$, i.e. standard deviation or variance proportional to expected value.)

Lemma 8.2: Under those distribution assumptions (above), it holds:

$$\text{SW}(x) = \mathbb{E}_{\text{v}}\left[\sum_{i=1}^{n} v_i(x_i)\right] > \text{SW}(\mathcal{A}) - \sqrt{2\log(nm)}\sum_{i=1}^{2m}\sigma_i$$

Proof. On any sample, the allocation of the a priori optimal mechanism, $\mathcal{A}$, has Expected Social Welfare equal to $\text{SW}(\mathcal{A}) = \mu_1 + \mu_2 + \cdots + \mu_m$, since for any product distribution F, $\mathcal{A}$ chooses the m marginal values with the highest expected value.

We want to bound the difference between the expected Social Welfare of that allocation x, on some random sample v (different from the one that determined the allocation x), and the expected Social Welfare of $\mathcal{A}$ on that same sample.

Let i be a specific marginal value that the VCG mechanism on the first sample chose to satisfy and let μ_i, σ_i be that marginal value's expected value and standard deviation, respectively. If $i \in [n]$ then that marginal value is also included in the allocation of $\mathcal{A}$, and it contributes nothing to the difference $\text{SW}(\mathcal{A}) - \text{SW}(x)$. So, the worst case scenario for the allocation x is if it chose marginal values that appeared good on the first sample, but are not part of the a priori optimal allocation, because they don't actually have that high expected value:

- As noted before, $\mathcal{A}$ chooses the m first marginal values in the sorted lists $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{mn}$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{mn}$.
- So, the highest standard deviation, out of all the marginal values that $\mathcal{A}$ does not choose is exactly σ_{m+1}.
- Even if all the $(n-1)m$ marginal values that $\mathcal{A}$ didn't choose were i.i.d. distributed $\sim G(\mu_{m+1}, \sigma_{m+1}^2)$ (or $SG(\mu_{m+1}, \sigma_{m+1}^2)$), their expected maximum would be at most $\mu_{m+1} + \sigma_{m+1}\sqrt{2\log((n-1)m)}$.
  (That is the maximum of n(m-1) random variables iid distributed according to $G(\mu_{m+1}, \sigma_{m+1}^2)$). Finding a reference for this was actually hard[4]! For the sake of completeness, we'll also include the proof in the end.
- The point is that on expectation, the most that a marginal value not in the allocation of $\mathcal{A}$ will "over-represent" itself on our first sample, compared to its actual expected value is $\sigma_{m+1}\sqrt{2\log((n-1)m)}$.
- Using the same logic, on expectation, the second-most that a marginal value not in the allocation of $\mathcal{A}$ will "over-represent" itself on the first sample, compared to its actual expected value is at most $\sigma_{m+2}\sqrt{2\log((n-1)m)}$.
- Using linearity of expectations and the fact that VCG can allocate at most m units, those m marginal values will on expectation be at most $\sqrt{2\log((n-1)m)}\sum_{i=1}^{m}\sigma_{m+i}$ higher on the first sample than their actual expected value.

In the same vein, using the fact that $\mathcal{A}$ also allocates at most m units, on expectation the most that the allocation of $\mathcal{A}$ could have been "under-represented" in the first sample is $\sqrt{2\log(nm)}\sum_{i=1}^{m}\sigma_i$.
(That would happen if for its i-th unit, all nm marginal values were i.i.d. Samples of $G(\mu_i, \sigma_i^2)$, and the marginal value that $\mathcal{A}$ actually chooses is the smallest of these samples.)

The allocation of VCG has higher expected Social Welfare than the allocation of the a priori optimal mechanism, $\mathcal{A}$:

$$\text{SW}(x) + \sqrt{2\log((n-1)m)}\sum_{i=1}^{m}\sigma_{m+i} \geq \text{SW}(\mathcal{A}) - \sqrt{2\log(nm)}\sum_{i=1}^{m}\sigma_i$$

$$\text{SW}(x) \geq \text{SW}(\mathcal{A}) - \sqrt{2\log((n-1)m)}\sum_{i=1}^{m}\sigma_{m+i} - \sqrt{2\log(nm)}\sum_{i=1}^{m}\sigma_i$$

$$\implies \text{SW}(x) > \text{SW}(\mathcal{A}) - \sqrt{2\log(nm)}\sum_{i=1}^{2m}\sigma_i$$

□

Lemmata 8.1 and 8.2 immediately imply the following theorem:

Theorem 8.3: In the setting of multi-unit auctions with submodular bidders, whose valuations follow MHR Gaussian/Sub-Gaussian distributions and the ordering of their standard deviations is the same as the ordering of their expected values, for the expected revenue of the Two Samples Mechanism, it holds:

$$\text{Rev(TSM)} \geq \frac{1}{4}\text{SW}(x) \geq \frac{1}{4}\left(\text{SW}(\mathcal{A}) - \sqrt{2\log(nm)}\sum_{i=1}^{2m}\sigma_i\right) \quad (11)$$

Combining the above theorem with the bound for SW($\mathcal{A}$) we showed on section 4, (C) results in the following Corollary:

Corollary 8.3.1: In the setting of multi-unit auctions with submodular bidders, whose valuations follow MHR Gaussian/Sub-Gaussian distributions and the ordering of their standard deviations is the same as the ordering of their expected values, for the expected revenue of the Two Samples Mechanism, it holds:

$$\text{Rev(TSM)} \geq \frac{1}{4}\text{SW(VCG)} - \frac{1}{4}\sqrt{2\log(nm)}\sum_{i=1}^{2m}\sigma_i \quad (12)$$
$$-\frac{1}{4}\sum_{i=1}^{m}\max\{\mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)} + \sigma_1\sqrt{2\log m}, 0\}$$

B. Regular Distributions

We examine the following setting:

- n bidders
- multi-unit auction with m units available for sale
- each bidder i is submodular: His "marginal value" for clinching additional units weakly decreases with the number of units he has clinched, i.e.: $v_{i,1} \geq v_{i,2} \geq \cdots \geq v_{i,m}$, where $v_{i,j}$: bidder i's value for clinching his j-th unit, provided he has already clinched j-1 units.
- All marginal values, $v_{i,j}$, follow a distribution with probability density function $f_{i,j}(\cdot)$.
- All those distributions are regular.

Lemma 8.4: Let $x = (x_1, x_2, \ldots, x_n)$ be the allocation that the Two Samples Mechanism determined on the first sample. Then, if all the bidders' distributions are regular, the expected revenue of the Two Samples Mechanism is at least a $\frac{1}{2}$-fraction of the expected revenue of posting to each bidder i the optimal price r_i^* for a bundle of x_i units.

Proof. Fix an arbitrary bidder i. Let F_i be the regular distribution of bidder i's value for receiving x_i units and r_i^* the optimal reserve price for that bundle of x_i units. Let $\widehat{R}_i(p) = p(1 - F_i(p))$ denote bidder i's revenue function for x_i units in value space. The key observation is that the price r_i that the Two Samples Mechanism will set for that bundle of units offered to bidder i is determined by the second sample, thus it is a random Sample from the distribution F_i. Applying [8, Lemma 3.6] for $t = 0$, we have:

$$\mathbb{E}_{v_i(x_i)}\left[\widehat{R}_i(\max\{v_i(x_i), 0\})\right] \geq \frac{1}{2}\cdot\widehat{R}(\max\{r_i^*, 0\}) \implies$$
$$\mathbb{E}_{v_i(x_i)}\left[\widehat{R}_i(v_i(x_i))\right] \geq \frac{1}{2}\cdot\widehat{R}(r_i^*)$$

Summing up over all bidders and using linearity of expectations:

$$\sum_{i=1}^{n}\mathbb{E}_{v_i(x_i)}\left[\widehat{R}_i(v_i(x_i))\right] \geq \frac{1}{2}\cdot\sum_{i=1}^{n}\widehat{R}(r_i^*) \implies$$
$$\mathbb{E}_{\text{v}\sim F}\left[\sum_{i=1}^{n}\widehat{R}_i(v_i(x_i))\right] \geq \frac{1}{2}\cdot\sum_{i=1}^{n}\widehat{R}(r_i^*)$$

□

Remark: It is easy to show that the above bound is tight. To see this, consider a regular valuation distribution F whose revenue function in probability space is essentially a triangle. For example, the distribution $F(v) = 1 - \frac{1}{v+1}$ on [0,H) and F(H) = 1 as $H \to \infty$.

Lemma 8.5: Let $x = (x_1, x_2, \ldots, x_n)$ be the allocation determined by the Two Samples Mechanism on the first sample. Let $B_i, \widetilde{\sigma_i}^2$ be the expected value and variance of F_i, bidder i's valuation distribution for x_i units. Then, if all the valuation distributions (all the F_i's) are regular, the following bound holds:

$$\text{Rev(TSM)} \geq \frac{1}{2}\cdot\text{SW}(x) - \frac{1}{2}\cdot\sum_{i=1}^{n}(3B_i)^{1/3}\widetilde{\sigma_i}^{2/3} \quad (13)$$

where Rev(TSM) is the expected revenue of the Two Samples Mechanism and
SW(x) is the expected Social Welfare of the allocation x on a random valuation profile, different than the one that induced the allocation.

Proof. The proof follows immediately from Lemma 7.4, [3, Corollary 2] and linearity of expectations:

$$\begin{aligned}\mathbb{E}_{\mathrm{V}}\left[\sum_{i=1}^{n}\widehat{R}_i(v_i(x_i))\right] &\geq \frac{1}{2}\cdot\sum_{i=1}^{n}\widehat{R}(r_i^*)\\ &= \frac{1}{2}\cdot\sum_{i=1}^{n}\left[\mathbb{E}_{v_i(x_i)\sim F_i}[v_i(x_i)] - S_i\right]\\ &= \frac{1}{2}\cdot\sum_{i=1}^{n}\mathbb{E}_{v_i(x_i)\sim F_i}[v_i(x_i)] - \frac{1}{2}\cdot\sum_{i=1}^{n}S_i\\ &\geq \frac{1}{2}\cdot\sum_{i=1}^{n}B_i - \frac{1}{2}\cdot\sum_{i=1}^{n}(3B_i)^{1/3}\widetilde{\sigma}_i^{\,2/3} \implies\end{aligned}$$

$$\mathrm{Rev(TSM)} \geq \frac{1}{2}\cdot\mathrm{SW}(x) - \frac{1}{2}\cdot\sum_{i=1}^{n}(3B_i)^{1/3}\widehat{\sigma}_i^{\,2/3}$$

Important Note: The part of section B below this point is **flawed:** That set of assumptions only holds for MHR distributions, since any Gaussian/Sub-Gaussian distribution is MHR. Thus, we can not claim that the final result of section B holds for regular distributions. We are working on correcting this point. The analysis is kept here, because it is used again in subsection C.

Up to this point, we have shown that the Expected Revenue of the Two Sample Mechanism is at least a $\frac{1}{2}$-fraction of the Expected Social Welfare of the allocation x, minus an additive term that is sublinear on the bidders' expected value and variance, for the number of units that they are allocated on x. Similarly to the case of MHR distributions, now, we need to show how close the Expected Social Welfare of x is to the Expected Social Welfare of the VCG mechanism. For this comparison, we'll make some additional assumptions about the distribution. These assumptions are the exact same that we used in our sample complexity bounds, and in the few samples case with regular distributions. The reason we are making these assumptions separately is to leave open the possibility of extending our results for different sets of distribution assumptions.

This set of additional assumptions is:

1) All marginal values are drawn exclusively from Gaussian and Sub-Gaussian distributions.
2) Let $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{mn}$ be their expected values, in sorted order.
3) Finally, we assume that all standard deviations/Sub-Gaussian parameters can be sorted in the same relative way, $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_{mn}$ where (μ_i, σ_i) correspond to the same distribution.
   (This is a weaker assumption that $\sigma_i \sim \mu_i$ or $\sigma_i^2 \sim \mu_i$, i.e. standard deviation or variance proportional to expected value.)

As we showed in Subsection 8.A, in this setting Lemma 8.2 holds. Combining it with Lemma 8.5, theorem 8.6 immediately follows:

Theorem 8.6: In the setting of multi-unit auctions with submodular bidders, whose valuations follow regular, Gaussian/Sub-Gaussian distributions and the ordering of their standard deviations/Sub-Gaussian parameters is the same as the ordering of their expected values, for the expected revenue of the Two Samples Mechanism, we have:

$$\begin{aligned}\mathrm{Rev(TSM)} &\geq \frac{1}{2}\cdot\mathrm{SW}(x) - \frac{1}{2}\cdot\sum_{i=1}^{n}(3B_i)^{1/3}\sigma_i^{2/3}\\ &\geq \frac{1}{2}\left(\mathrm{SW}(\mathcal{A}) - \sqrt{2\log(nm)}\sum_{i=1}^{2m}\sigma_i\right)\\ &\quad - \frac{1}{2}\cdot\sum_{i=1}^{n}(3B_i)^{1/3}\widetilde{\sigma}_i^{\,2/3}\end{aligned} \tag{14}$$

Combining the above theorem with the bound for $\mathrm{SW}(\mathcal{A})$ we showed on ... results in the following Corollary:

Corollary 8.6.1: In the setting of multi-unit auctions with submodular bidders, whose valuations follow regular Gaussian and Sub-Gaussian distributions and the ordering of their standard deviations is the same as the ordering of their expected values, for the expected revenue of the Two Samples Mechanism, we have:

$$\begin{aligned}\mathrm{Rev(TSM)} \geq{}& \frac{1}{2}\mathrm{SW(VCG)} - \frac{1}{2}\sqrt{2\log(nm)}\sum_{i=1}^{2m}\sigma_i\\ &- \frac{1}{2}\cdot\sum_{i=1}^{n}(3B_i)^{1/3}\widetilde{\sigma}_i^{\,2/3}\\ &- \frac{1}{2}\sum_{i=1}^{m}\max\{\mu_{m+i} - \mu_{m+1-i} + \sigma_{m+i}\sqrt{2\log((n-1)m)}\\ &+ \sigma_1\sqrt{2\log m}, 0\}\end{aligned} \tag{15}$$

C. Extending the previous result for MHR distributions

The same analysis of the previous section can also be applied to MHR distributions, guaranteeing the same approximation. However, in the case of MHR distributions, with a very small tweak on the Two Samples Mechanism, we can increase the approximation ratio.

Algorithm 7 The Two Samples MHR Mechanism (TSM′)

1: Collect 2 samples of all bidders' marginal valuations.
2: Run VCG on the first sample. Let $x = (x_1, x_2, \ldots, x_n)$ be the allocation of that VCG.
3: In the auction, offer to each bidder i a bundle of size x_i, at a price equal to that bidder i's value for x_i units in the second sample multiplied by 0.85.

Essentially, we use one sample to determine the size of the bundle offered to each bidder, and another to determine the price of each bundle.

From [10, Theorem 5.1], for MHR distributions setting the reserve price for a bundle equal to 0.85 of a random sample of that distribution achieves expected revenue greater or equal to 0.589 of the expected revenue of the optimal reserve price. Replacing Lemma 7.4 of the previous subsection with this result, and continuing the rest of the analysis in the exact same way yields the following corollary:

Corollary 8.6.2: In the setting of multi-unit auctions with submodular bidders, whose valuations follow MHR, (Gaussian/)Sub-Gaussian distributions and the ordering of their standard deviations is the same as the ordering of their expected values, for the expected revenue of the Two Samples Mechanism (with modified reserves), we have:

$$\begin{aligned}\text{Rev(TSM)} \geq 0.589 \cdot \text{SW(VCG)} - 0.589 \cdot \sqrt{2\log(nm)} \sum_{i=1}^{2m} \sigma_i \\ - 0.589 \cdot \sum_{i=1}^{n} (3B_i)^{1/3} \widetilde{\sigma_i}^{2/3} \\ - 0.589 \cdot \sum_{i=1}^{m} \max\{\mu_{m+i} - \mu_{m+1-i} + \\ + \sigma_{m+i}\sqrt{2\log((n-1)m)} + \sigma_1 \sqrt{2\log m}, 0\}\end{aligned} \tag{16}$$

IX. References

1) Morgenstern,Roughgarden - The Pseudo-Dimension of Near-Optimal Auctions
2) Balcan - A General Theory of Sample Complexity for Multi-Item Profit Maximization
3) Medina, Vassilvitskii -Revenue Optimization with Approximate Bid Predictions
4) Pollard - Convergence of Stochastic Processes
5) Gkatzelis, Markakis, Roughgarden - Deferred-Acceptance Auctions for Multiple Levels of Service
6) Cai, Daskalakis - Extreme-value theorems for optimal multidimensional pricing
7) Hartline, Roughgarden - Simple Versus Optimal Mechanisms
8) Dhangwatnotai1, Roughgarden - Revenue Maximization with a Single Sample
9) Pascal Massart - Concentration Inequalities and Model Selection, Springer
15) P. Milgrom. 2004. Putting Auction Theory to Work. Cambridge University Press
10) Huang, Roughgarden - Making the Most of your Samples
11) P. Milgrom and I. Segal. 2014. Deferred-acceptance auctions and radio spectrum reallocation. In ACM Conference on Economics and Computation
12) FCC. 2017. Broadcast Incentive Auction. Corresponding Link
13) S. Li. 2015. Obviously Strategy-Proof Mechanisms. In SSRN. Working Paper No. 2560028.
14) P. Milgrom and I. Segal. 2015. Deferred-Acceptance Auctions and Radio Spectrum Reallocation. Available from http://web.stanford.edu/ isegal/heuristic.pdf
16) A. Ockenfels, D. H. Reiley, and A. Sadrieh. 2006. Economics and Information Systems. Handbooks in Information Systems, Vol. 1. Elsevier, Chapter 12. Online Auctions, 571–628.
17) R. Cole and T. Roughgarden. The Sample Complexity of revenue maximization. In ACM Conference on Economics and Computation 2014
18) A. Medina and M. Mohri. Learning theory and algorithms for revenue optimization in second price auctions with reserve. In Proceedings of the International Conference on Machine Learning (ICML).
19) J. Morgenstern and T. Roughgarden. Learning simple auctions. In Proceedings of the Conference on Learning Theory (COLT).
20) C. Gresse and J. Gajewski. A Survey of the European IPO market. ECMI paper 2006
21) Kjell G. Nyborg, Kristian Rydqvist and Suresh M. Sundaresan.Bidder Behavior in Multiunit Auctions: Evidence from Swedish Treasury Auctions
22) He Huang, Yu-e Sun, Kai Xing, Hongli Xu, Xueyong Xu, Liusheng Huang.Truthful Multi-unit Double Auction for Spectrum Allocation in Wireless Communications. International Conference on Wireless Algoirthms, Systems, and Applications (WASA 2012)
23) P. Dütting, V. Gkatzelis, and T. Roughgarden. 2014. The performance of deferred-acceptance auctions. In ACM Conference on Economics and Computation, EC ’14.