Learning Within an Instance for Designing High-Revenue Combinatorial Auctions

Maria-Florina Balcan[1] , **Siddharth Prasad**[1] and **Tuomas Sandholm**[1,2,3,4]
[1]School of Computer Science, Carnegie Mellon University
[2]Optimized Markets, Inc.
[3]Strategic Machine, Inc.
[4]Strategy Robot, Inc.
{ninamf, sprasad2, sandholm}@cs.cmu.edu

Abstract

We develop a new framework for designing truthful, high-revenue (combinatorial) auctions for limited supply. Our mechanism *learns within an instance*. It generalizes and improves over previously-studied random-sampling mechanisms. It first samples a *participatory group* of bidders, then samples several *learning groups* of bidders from the remaining pool of bidders, learns a high-revenue auction from the learning groups, and finally runs that auction on the participatory group. Previous work on random-sampling mechanisms focused primarily on unlimited supply. Limited supply poses additional significant technical challenges, since allocations of items to bidders must be feasible. We prove guarantees on the performance of our mechanism based on a market-shrinkage term and a new complexity measure we coin *partition discrepancy*. Partition discrepancy simultaneously measures the intrinsic complexity of the mechanism class and the uniformity of the set of bidders. We then introduce new auction classes that can be parameterized in a way that does not depend on the number of bidders participating, and prove strong guarantees for these classes. We show how our mechanism can be implemented efficiently by leveraging practically-efficient routines for solving winner determination. Finally, we show how to use structural revenue maximization to decide what auction class to use with our framework when there is a constraint on the number of learning groups.

1 Introduction

In a (limited-supply) *combinatorial auction*, a seller has m indivisible items to allocate among a set S of n bidders. Combinatorial auctions have various real-world applications. Two examples include auctions for allocating licenses for bands of the electromagnetic spectrum and sourcing auctions for supply chain management. The design of truthful, high-revenue combinatorial auctions is a central problem in mechanism design. A comprehensive account of combinatorial auctions may be found in Cramton et al. [2006].

A common strategy for designing truthful, high-revenue auctions when there is an *unlimited supply* of each good has been to use a random-sampling mechanism. A random-sampling mechanism splits the bidders into two groups, and applies the optimal auction for each group to the other group (thereby achieving truthfulness, since the auction run on any bidder's group is independent of her reported valuation). In unlimited-supply settings, random-sampling mechanisms satisfy strong guarantees [Goldberg *et al.*, 2001; Balcan *et al.*, 2005; Alaei *et al.*, 2009].

However, until now, there has been no unified, general-purpose method of adapting the random-sampling approach to analyze the limited-supply setting. Limited supply poses additional significant technical challenges, since allocations of items to bidders must be feasible. For example, random-sampling with any mechanism class that allows bidders to purchase according to their demand functions would violate supply constraints. Most adaptations of random-sampling to limited supply deal with feasibility issues in complicated ways, for example, by constructing intricate revenue benchmarks to limit the number of buyers who can make a purchase [Balcan *et al.*, 2007], or by placing combinatorial constraints on the environment [Devanur and Hartline, 2009; Devanur *et al.*, 2015].

In this paper we circumvent these issues by applying auction formats that generalize the classical Vickrey-Clarke-Groves (VCG) auction [Vickrey, 1961; Clarke, 1971; Groves, 1973] to sell all m items to a random group of participatory bidders. These auctions prescribe feasible allocations and payments (and are incentive compatible). Several parameterized generalizations of the VCG auction have been studied with the aim of increasing revenue by introducing weights to favor certain bidders or allocations. Examples include affine-maximizer auctions (AMAs) [Roberts, 1979], virtual-valuations combinatorial auctions (VVCAs) [Likhodedov and Sandholm, 2004; Likhodedov and Sandholm, 2005; Sandholm and Likhodedov, 2015], λ-auctions [Jehiel *et al.*, 2007], mixed-bundling auctions with reserve prices [Tang and Sandholm, 2012], and mixed-bundling auctions [Jehiel *et al.*, 2007]. However, little is known when it comes to formal approximation guarantees for these auction classes.

A direct adaptation of vanilla random sampling can do poorly when the auction class is rich. Suppose we randomly partition the set of bidders into two groups S^1 and S^2, and apply the optimal mechanism for S^1 to S^2. Consider learning a second-price auction with a reserve in the case of selling a single item. Suppose there is one bidder who values the item at 10 and the remaining buyers' values are in $[0, 9]$. The high bidder is in S^1 with probability $1/2$. So with probability $1/2$, the optimal reserve price for S^1 is 10, and the revenue obtained from S^2 is 0. More generally, since we study large parameterized auction classes, the optimal auction for S^1 potentially overfits to a small number of bidders. Another adaption along the lines of vanilla random sampling to prevent overfitting would be to partition the set S of bidders into N groups, use the first $N-1$ groups to learn a high-revenue auction, and then apply that auction to the Nth group. The issue with this approach is that generalization guarantees would require N large. Thus the final mechanism only sells items to a tiny fraction of bidders, incurring a large revenue loss.

Our main *learning-within-an-instance (LWI) mechanism* alleviates these issues by randomly drawing a set of participatory bidders S_{par}, and then sampling several proportionally-sized learning groups from $S_{lrn} := S \setminus S_{par}$ to learn an auction that is close-to-optimal *in expectation* for a random learning group. Our approach is a form of *automated mechanism design* [Conitzer and Sandholm, 2002; Sandholm, 2003].

1.1 Setup and the Main Mechanism

In our model, the seller has m indivisible items to allocate among a set S of n bidders/buyers. Each buyer is described by her valuation function $v_i : 2^{\{1,\ldots,m\}} \to \mathbb{R}_{\geq 0}$ over bundles of the m items. (We implicitly assume that each buyer's value for getting the empty bundle is zero.) We do not assume that $b \subseteq b'$ implies $v(b) \leq v(b')$ (a common assumption called free disposal). For an allocation α, $v_i(\alpha)$ denotes the value buyer i assigns to the bundle she receives according to α (we assume that buyers valuations are independent of what other buyers' receive). For an allocation α, $W(\alpha) = \sum_{i=1}^{n} v_i(\alpha)$ denotes the welfare of α, and $W_{-i}(\alpha) = \sum_{j \neq i} v_j(\alpha)$ denotes the welfare of α when bidder i is absent. For a set of bidders S, $W(S) = \max_\alpha W(\alpha)$ denotes the welfare of an *efficient* allocation, that is, an allocation that maximizes welfare. The VCG auction uses the efficient allocation α^*, and bidder i pays $\max_\alpha W_{-i}(\alpha) - W_{-i}(\alpha^*)$. The auctions we study in this paper are parameterized generalizations of the VCG auction that modify the welfare function by applying boosts to specific allocations with the aim of increasing revenue. For an auction M and a set of bidders $S' \subseteq S$, we denote by $\mathsf{Rev}_M(S')$ the sum of the payments made by bidders in S' when the seller runs M among bidders in S'. We write $S' \sim_p S$ to denote a subset S' being sampled from S by including each bidder in S' independently with probability p.

We now present the main mechanism of this paper.

Learning-within-an-instance mechanism (LWI) Parameters: p, q, N

1. Draw a group of participatory buyers $S_{par} \sim_p S$.
2. Draw learning groups of buyers $S_1, \ldots, S_N \sim_q S \setminus S_{par}$.
3. Find the mechanism $\widehat{M} \in \mathcal{M}$ that maximizes empirical revenue $\frac{1}{N}\sum_{t=1}^{N} \mathsf{Rev}_M(S_t)$ over the learning groups.
4. Apply mechanism $\widehat{M}$ to S_{par}.

When $\mathcal{M}$ is a class of incentive-compatible mechanisms, LWI is incentive-compatible since $\widehat{M}$ does not depend on the valuations of the bidders in S_{par}.

1.2 Summary of the Contributions of this Paper

In section 2 we provide the main guarantees satisfied by our LWI framework. The guarantees are derived using learning-theoretic techniques. Informally, they provide (high probability) lower bounds on the performance of LWI of the form $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)(L_{\mathcal{M}} - \varepsilon_{\mathcal{M}}(N, \delta)) - \tau_{\mathcal{M}}$, where $L_{\mathcal{M}}$ measures the revenue loss incurred by allocating items only to participatory bidders, $\varepsilon_{\mathcal{M}}$ is a standard learning-theoretic error term that depends on the intrinsic complexity of $\mathcal{M}$, and $\tau_{\mathcal{M}}$ is an additional error term we coin *partition discrepancy*. Partition discrepancy is also a measure of the intrinsic complexity of $\mathcal{M}$, but is simultaneously a measure of the level of uniformity in the set S of bidders. We provide examples and a general bound to illustrate properties of partition discrepancy.

In section 3 we introduce a new class of auctions called *bundling-boosted auctions*. These auctions are parameterized in a way that does not depend on the number of bidders who participate in the auction (unlike most previous generalizations of the VCG auction). We prove bounds on the intrinsic complexity of bundling-boosted auctions (and a few other natural subclasses of auctions) that have no dependence on the number of bidders. We show that under certain conditions LWI on the class of bundling-boosted auctions yields an $(O(p) - \varepsilon)$-approximation with high probability.

In section 4 we show how our learning-within-an-instance mechanism can be implemented in a sample and computationally efficient manner for bundling-VCG auctions and sparse bundling-boosted auctions by leveraging practically efficient routines for solving winner determination.

In section 5 we show how to use structural revenue maximization to decide what auction class to use with LWI when there is a constraint on the number of learning groups.

1.3 Additional Related Research

There have been various alternate approaches to revenue maximization for limited supply. Balcan, Blum, and Mansour [2008] obtain a $O(2^{\sqrt{\log m \log\log m}})$-approximation for bidders with subadditive valuations, which was improved to a $O(\log^2 m)$-approximation by Chakraborty, Huang, and Khanna [2013]. Both these works studied item-pricing mechanisms. Sandholm and Likhodedov [2005; 2015] obtain a $(2 + 2\log(h/l))$-approximation when bidders have additive valuations, where l and h are lower and upper bounds on the valuation of any bidder for any item. Our results significantly improve upon these existing results in various situations. For example, for $W(S)$ sufficiently large, we prove that our LWI mechanism run on the class of bundling-boosted auctions yields an $(O(p) - \varepsilon)$-approximation. In addition, previous approximations are on expected revenue, while we give the much stronger guarantee of *high-probability* revenue approximation. Furthermore, our results do not require restrictions on valuation functions, giving them very broad applicability.

A recent line of work studies learning revenue-maximizing auctions for limited supply *across instances* [Mohri and Medina, 2014; Morgenstern and Roughgarden, 2015; Balcan *et* *al.*, 2018]. These works laid down the framework for understanding learning-theoretic quantities related to auctions in order to prove generalization guarantees. Our paper studies the significantly tougher and unsolved problem of learning from a single instance for limited supply. We extend the techniques of Balcan et al. [2005] (that can be viewed as learning within an instance for unlimited supply) and show that learning theory combined with the power of parameterized auctions provides a way to meaningfully learn within an instance in the more challenging setting of limited supply.

2 Main Guarantees of our Framework

In this section we present the main guarantees satisfied by LWI in terms of structural properties of the auction class and the set of bidders. Our guarantees are in terms of partition discrepancy, delineability, and the following quantity that controls the revenue loss incurred by selling only to bidders in S_{par}. For $S' \subseteq S$, let $\mathrm{OPT}_{\mathcal{M}}(S') = \sup_{M\in\mathcal{M}} \mathsf{Rev}_M(S')$ and let $L_{\mathcal{M}}(S') = \mathrm{OPT}_{\mathcal{M}}(S')/W(S)$.

For a given participatory set of bidders S_{par}, *partition discrepancy* measures the worst-case deviation in an auction class between the revenue on S_{par} versus the expected revenue on a set of bidders sampled from $S\setminus S_{par}$. For $0<q<1$ and $S_{par}\subset S$, partition discrepancy is defined as

$$\tau_{\mathcal{M}}(q, S_{par}) = \sup_{M\in\mathcal{M}} \left|\mathsf{Rev}_M(S_{par}) - \mathop{\mathbf{E}}_{S_0\sim_q S\setminus S_{par}}[\mathsf{Rev}_M(S_0)]\right|.$$

Partition discrepancy is a measure of both the intrinsic complexity of the class $\mathcal{M}$ and the amount of uniformity in the set S of bidders. We now present general guarantees for LWI in terms of partition discrepancy (the full derivations are in the appendix). The guarantees follow from uniform convergence results, and depend on the expected Rademacher complexity $R_{\mathcal{M}}(N; S\setminus S_{par})$ of $\mathcal{M}$ with respect to $S\setminus S_{par}$ and the pseudodimension $\mathrm{Pdim}(\mathcal{M})$ of $\mathcal{M}$. We provide definitions and some standard results from learning theory that we use in our proofs in the appendix. $\widehat{M}$ denotes the empirical-revenue-maximizing mechanism used by LWI.

Theorem 1. *Let S_{par} denote the participatory set of bidders chosen by a run of LWI. Then, with probability $\geq 1-2\delta$ over the draw of $S_1,\dots,S_N\sim_q S\setminus S_{par}$, (a) $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)\Big(L_{\mathcal{M}}(S_{par}) - 4R_{\mathcal{M}}(N; S\setminus S_{par}) - \sqrt{2\ln(1/\delta)/N}\Big) - 2\tau_{\mathcal{M}}(q, S_{par})$ and (b) $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)\Big(L_{\mathcal{M}}(S_{par}) - 240\sqrt{\mathrm{Pdim}(\mathcal{M})/N} - \sqrt{2\ln(1/\delta)/N}\Big) - 2\tau_{\mathcal{M}}(q, S_{par})$.*

Proof sketch. Uniform convergence results relate the empirical revenue of $\widehat{M}$ on the learning groups to the optimal expected revenue on a random learning group. Partition discrepancy ties both these quantities to revenue on S_{par}. □

If $\mathcal{M}$ has finite pseudodimension (this is not necessarily the case if we only have a bound on Rademacher complexity), we can give an equivalent sample-complexity version of the guarantee. Let $N(\varepsilon,\delta,\mathrm{Pdim}(\mathcal{M})) = 480^2\,\mathrm{Pdim}(\mathcal{M})\ln\left(\frac{1}{\delta}\right)/\varepsilon^2$.

Corollary 1. *Let S_{par} denote the participatory set of bidders chosen by a run of LWI with parameters p,q,N, where $N\geq N(\varepsilon,\delta,\mathrm{Pdim}(\mathcal{M}))$. Then, with probability $\geq 1-2\delta$ over the draw of $S_1,\dots,S_N\sim_q S\setminus S_{par}$, $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)(L_{\mathcal{M}}(S_{par})-\varepsilon) - 2\tau_{\mathcal{M}}(q, S_{par})$.*

To understand the pseudodimension of various mechanism classes, Balcan et al. [2018] introduced the notion of *delineability*. A class of mechanisms $\mathcal{M}$ is (d,h)-*delineable* if (1) every $M\in\mathcal{M}$ can be parameterized by a vector $\theta\in\mathbb{R}^d$, and (2) for every set S of bidder valuations, there are at most h hyperplanes partitioning $\mathbb{R}^d$ such that $\mathsf{Rev}_S(\theta) := \mathsf{Rev}_\theta(S)$ is linear in θ over each connected component of $\mathbb{R}^d$ determined by the hyperplanes. The way we have stated delineability requires h to be independent of the number of bidders in S. We include an analysis of the case where h is allowed to be a function of n in the appendix. The following example illustrates delineability in a simple case. Balcan et al. [2018] provide more examples and a more detailed discussion.

Example 1 (Second-price auctions with a reserve price)**.** *The class of second-price auctions with reserve prices for selling a single item is $(1,2)$-delineable. Indeed, if v_1 and v_2 are highest and second-highest values for the item, respectively, then for $r<v_2$ the revenue of a second-price auction with reserve r is v_2, for $v_2\leq r\leq v_1$ it is r, and for $r>v_1$ it is 0.*

Rademacher complexity, pseudodimension, and delineability are connected through the following relations: $R_{\mathcal{M}}(N; S_{lrn}) \leq 60W(S)\sqrt{\mathrm{Pdim}(\mathcal{M})/N}$ [Dudley, 1987] and if $\mathcal{M}$ is (d,h)-delineable, $\mathrm{Pdim}(\mathcal{M})\leq 9d\ln(4dh)$ [Balcan *et al.*, 2018].

We present our main guarantee in terms of delineability:

Theorem 2. *Suppose $\mathcal{M}$ is (d,h)-delineable. Let S_{par} denote the participatory set of bidders chosen by a run of LWI with parameters p,q,N, where $N\geq N(\varepsilon,\delta,9d\ln(4dh))$. Then, with probability $\geq 1-2\delta$ over the draw of $S_1,\dots,S_N\sim_q S\setminus S_{par}$, $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)(L_{\mathcal{M}}(S_{par})-\varepsilon) - 2\tau_{\mathcal{M}}(q, S_{par})$.*

We provide analogous guarantees for mechanism classes that satisfy a version of delineability that is dependent on the number of bidders in the appendix.

2.1 Partition Discrepancy

In this section we develop a further understanding of partition discrepancy. We first provide two examples illustrating structural properties of partition discrepancy. We then provide a general-purpose high-probability bound on partition discrepancy based on pseudodimension of the mechanism class.

The first example relates the failure of vanilla random sampling to large partition discrepancy using the scenario given in the introduction. We show how LWI alleviates that issue.

Example 2 (LWI versus random sampling)**.** *Consider the example from the introduction where a single item is for sale and $\mathcal{M}$ is the class of second-price auctions with reserve. There is one bidder with value 10, and all remaining bidders' values are in $[0,9]$. Suppose LWI is run with parameters $p=1/2, q=1$ (which corresponds to vanilla random sampling). Then, for any participatory set S_{par}, $\tau_{\mathcal{M}}(1,S_{par}) = 10 = W(S)$, achieved by setting a reserve price of 10. If instead LWI was run with parameters $p=q<1$, the high bidder is in $S\setminus S_{par}$ with probability $1-p$, and in this case*

$\tau_{\mathcal{M}}(q, S_{par}) = 10q$ *If, for example,* $p = q = 1/20$, *this is a small additive loss in the overall revenue guarantee.*

The next example involves replica economies, where the set of bidders is composed of several copies of a ground set of bidders. Replica economies have been studied extensively in economics (and recently from an algorithmic viewpoint) in the context of convergence to equilibria [Debreu and Scarf, 1963; Aumann, 1964; Barman and Echenique, 2020].

Example 3 (Replica economies)**.** *Suppose* $S_0 = \{v_1, v_2, v_3\}$, *and* S *consists of* n_0 *replicas of* S_0. *Let* $\mathcal{M}$ *be a population-size-independent auction class. We show in the appendix that if* n_0 *is sufficiently large,* $\tau_{\mathcal{M}}(1, S_{par}) = 0$ *with high probability over the draw of* $S_{par} \sim_{1/2} S$.

We now present a general bound on partition discrepancy in terms of the learning-theoretic complexity of $\mathcal{M}$ when LWI is run with parameters $p = 1/3$ and $q = 1/2$. For each bidder i, let $\tilde{v}_i = \max_{|S'| \geq n/3 - 5\sqrt{n}} \sup_{M \in \mathcal{M}} |\mathsf{Rev}_M(S' \cup \{i\}) - \mathsf{Rev}_M(S')|$ and let $\tilde{v} = (\tilde{v}_1, \ldots, \tilde{v}_n) \in \mathbb{R}^n$. These terms measure how sensitive the mechanism class is to the addition of a single bidder to an already large set of bidders. In the following results on partition discrepancy, we condition on the (probability $\geq 1 - e^{-25}$) event that $|S_{par}| \geq n/3 - 5\sqrt{n}$.

Theorem 3. *With probability* $\geq 1 - \delta$ *over the draw of* $S_{par} \sim_{1/3} S$, $\tau_{\mathcal{M}}(1/2, S_{par}) \leq ||\tilde{v}||_2 \sqrt{2n \operatorname{Pdim}(\mathcal{M}) \ln \frac{4e^2 \operatorname{Pdim}(\mathcal{M}) W(S)}{\delta}}$.

Proof sketch. We bound τ_M for a single mechanism M using concentration bounds. We then apply a union bound over a learning-theoretic cover of $\mathcal{M}$. Classical learning-theory results bound the cover size in terms of $\operatorname{Pdim}(\mathcal{M})$. ☐

Combined with Corollary 1, we have:

Theorem 4. *Run LWI with parameters* N, $p = 1/3$, $q = 1/2$, *where* $N \geq N(\varepsilon, \delta, \operatorname{Pdim}(\mathcal{M}))$. *Then, with probability* $\geq 1 - 3\delta$, $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)(L_{\mathcal{M}}(S_{par}) - \varepsilon) - 2||\tilde{v}||_2 \sqrt{2n \operatorname{Pdim}(\mathcal{M}) \ln \frac{4e^2 \operatorname{Pdim}(\mathcal{M}) W(S)}{\delta}}$.

When $W(S)$ is sufficiently large, we can condense the bound on partition discrepancy to contribute at most an ε loss.

Corollary 2. *Run LWI with parameters* N, $p = 1/3$, $q = 1/2$, *where* $N \geq N(\varepsilon, \delta, \operatorname{Pdim}(\mathcal{M}))$. *If* $W(S)^2 - \frac{8n||\tilde{v}||_2^2 \operatorname{Pdim}(\mathcal{M})}{\varepsilon^2} \ln(W(S)) \geq \frac{8n||\tilde{v}||_2^2 \operatorname{Pdim}(\mathcal{M})}{\varepsilon^2} \ln\left(\frac{4e^2 \operatorname{Pdim}(\mathcal{M})}{\delta}\right)$, $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)(L_{\mathcal{M}}(S_{par}) - 2\varepsilon)$ *with probability* $\geq 1 - 3\delta$.

Remark. We emphasize that small partition discrepancy (for example, stipulating that $\tau_{\mathcal{M}}$ is a fixed constant) should be viewed as a uniformity condition on the set of bidders. Theorem 3 provides just one way of understanding partition discrepancy by relating it to learning-theoretic quantities.

3 Population-Size-Independent Auctions

In this section we instantiate our main guarantee for specific mechanism classes $\mathcal{M}$ to obtain more concrete revenue approximations. The following is a naïve lower bound on $L_{\mathcal{M}}(S_{par})$ for auction classes that can run a second-price auction on the grand bundle $\{1, \ldots, m\}$ with a reserve price.

Proposition 1. *Let* $v_1 \geq \cdots \geq v_n$ *denote the valuations of each bidder in* S *on the grand bundle. For any* $0 < \alpha \leq 1$ *such that* αn *is an integer, any mechanism class* $\mathcal{M}$ *containing the second-price auction on the grand bundle with reserve price* r *for every* r *satisfies* $L_{\mathcal{M}}(S_{par}) \geq \frac{v_{\alpha n}}{W(S)}$ *with probability* $\geq 1 - e^{-\alpha n p}$ *over the draw of* $S_{par} \sim_p S$.

However, any bidder's value for the grand bundle can be an arbitrarily bad approximation to $W(S)$. In the remainder of the paper we introduce some new auction classes and prove more fine-tuned approximations for those classes.

We now study a handful of *population-size-independent* auction classes, that is, auction classes that can be parameterized in a way that does not depend on the number of bidders. Traditional variants to the VCG auction including λ-auctions and AMAs specify boosts based on particular allocations and are thus not independent of the population size (and in particular cannot be used with LWI in a natural way since $S_1, \ldots, S_N, S_{par}$ can all vary in size). In contrast to these, our auctions specify boosts based on bundles and bundlings.

A bundling is a partition of items $\{1, \ldots, m\}$ into bundles. We say that an allocation *respects* a bundling if no two items in the same bundle are allocated to different buyers. For an allocation β, let $\mathsf{blg}(\beta)$ denote the finest bundling respected by β, that is, the bundling with the fewest number of bundles that β respects. For example, if β allocates items 1 and 3 to bidder 1, and the remaining items to bidder 2, $\mathsf{blg}(\beta) = \{\{1, 3\}, \{2, 4, \ldots, m\}\}$. Let Φ denote the collection of all bundlings. $|\Phi| < (0.792m/\ln(m+1))^m$ [Berend and Tassa, 2010]. We now introduce two new auction classes that can be viewed as population-size-independent analogues of λ-auctions and VVCAs, respectively.

The class of *bundling-boosted auctions* is the class auctions parameterized by real $|\Phi|$-dimensional vectors $\omega \in \mathbb{R}^{|\Phi|}$ that specify additive boosts $\omega(\phi)$ for each bundling $\phi \in \Phi$. The overall allocation α^* used by a bundling-boosted auction ω is chosen to maximize $W(\alpha) + \omega(\mathsf{blg}(\alpha))$, and bidder i pays $\max_\alpha (W_{-i}(\alpha) + \omega(\mathsf{blg}(\alpha))) - (W_{-i}(\alpha^*) - \omega(\mathsf{blg}(\alpha^*)))$. Equivalently, ω is the λ-auction with $\lambda(\alpha) = \omega(\mathsf{blg}(\alpha))$.

The class of *bundle-boosted auctions* is the class of auctions parameterized by real 2^m-dimensional vectors $\omega \in \mathbb{R}^{2^m}$ that specify additive boosts $\omega(b)$ for each bundle $b \subseteq \{1, \ldots, m\}$. The overall allocation α^* is chosen to maximize $W(\alpha) + \sum_{b \in \mathsf{blg}(\alpha)} \omega(b)$, and bidder i pays $\max_\alpha (W_{-i}(\alpha) + \sum_{b \in \mathsf{blg}(\alpha)} \omega(b)) - (W_{-i}(\alpha^*) + \sum_{b \in \mathsf{blg}(\alpha^*)} \omega(b))$. Equivalently, the class of bundle-boosted auctions is the subclass of VVCAs where the parameters are constant across bidders.

The class of bundling-VCG auctions [Kroer and Sandholm, 2015] consists of all ϕ-VCG auctions, where a ϕ-VCG auction runs VCG while treating each bundle in ϕ as an indivisible item. The class of bundling-VCG auctions is a subclass of the class of bundle-boosted auctions: the ϕ-VCG auction can be represented by the bundle-boosted auction with $\omega(b) = 0$ if b can be represented as a union of bundles from ϕ, and $\omega(b) = -\infty$ otherwise. The class of bundle-boosted auctions is a subclass of the class of bundling-boosted auctions: a bundle-boosted auction is a bundling-boosted auction with the restriction that $\omega(\phi) = \sum_{b \in \phi} \omega(b)$.

Since bundling-boosted and bundle-boosted auctions are

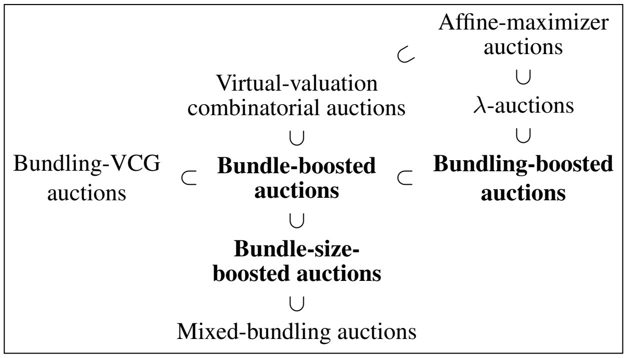

Figure 1: Containment relations between auction classes. New auction classes introduced in this paper are in boldface.

subclasses of λ-auctions, they are both delineable with $h(n) = (n+1)^{2m+1}$ due to Balcan et al. [2018] (see the appendix for a derivation of LWI guarantees when only weaker delineability parameters that depend on the number of bidders are known). The following is a much stronger delineability result that has no dependence on the number of bidders.

Theorem 5. *The class of bundling-boosted auctions is* $(|\Phi|, |\Phi|^2 + m|\Phi|^3)$*-delineable and the class of bundle-boosted auctions is* $(2^m, |\Phi|^2 + m|\Phi|^3)$*-delineable.*

Proof sketch. We prove that there are at at most $m|\Phi|$ bidders whose absence affects the allocations used by *any* bundling-boosted auction. This allows us to count the relevant hyperplanes delineating $\mathbb{R}^{|\Phi|}$ in a way that is independent of n. $\square$

Sandholm and Likhodedov [2005; 2015] (implicitly) study properties of the class of auctions parameterized by vectors $\omega \in \mathbb{R}^m$ that specify additive boosts depending on the size of the bundle. We call this class of auctions *bundle-size-boosted* auctions. Bundle-size-boosted auctions are a subclass of bundle-boosted auctions: the equivalent bundle-boosted auction satisfies $\omega(b) = \omega(|b|)$. For the class of bundle-size-boosted auctions, we can prove a stronger delineability result.

Theorem 6. *The class of bundle-size-boosted auctions is* $(m, me^{O(\sqrt{m})})$*-delineable.*

Figure 1 summarizes the containment relations between the various auction classes.

3.1 Guarantees for Bundling-Boosted Auctions

The class of bundling-boosted auctions is a rich class of auctions. If the efficient allocation when bidder i is absent also maximizes welfare when all bidders are present among all allocations respecting the same finest bundling, there is a bundling-boosted auction that extracts revenue equal to the welfare of the efficient allocation. More generally:

Theorem 7. *Given a set of S bidders, let β denote the efficient allocation and let β_{-i} denote the efficient allocation when bidder i is absent. Let* $\Delta_i(S) = \max_{\alpha:\mathsf{blg}(\alpha)=\mathsf{blg}(\beta_{-i})} W(\alpha) - W(\beta_{-i})$. *There exists a bundling-boosted auction with revenue* $W(S) - \sum_i \Delta_i(S)$.

Proof sketch. Let $\phi = \mathsf{blg}(\beta)$ and let $\phi_{-i} = \mathsf{blg}(\beta_{-i})$. The bundling-boosted auction with $\omega(\phi) = 0, \omega(\phi_{-i}) = W(\beta) - W(\beta_{-i})$, and $\omega(\phi') = -\infty$ for all other bundlings $\phi' \in \Phi \setminus \{\phi, \phi_{-1}, \ldots, \phi_{-n}\}$ extracts the claimed revenue. $\square$

We give a simple example of bidder valuations that satisfy $\Delta_i(S') = 0$ for every i and every $S' \subseteq S$ in the appendix.

Concentration inequalities enable us to provide bounds on $L_{\mathcal{M}}(S_{par})$ for the class of bundling-boosted auctions. We have $\mathbf{E}_{S_{par}}[\mathrm{OPT}_{\mathcal{M}}(S_{par})] \geq \mathbf{E}_{S_{par}}[W(S_{par}) - \sum_i \Delta_i(S_{par})] \geq pW(S) - \sum_i \Delta_i(S_{par})$. If we run LWI with parameters p, q, N, we have (assuming for readability that $\Delta_i(S_{par}) = 0$ for all i) $L_{\mathcal{M}}(S_{par}) \geq (1-\eta)p$ with probability $\geq 1 - e^{-2\eta^2 p^2 W(S)^2/||\bar{v}||_2^2}$, where $\bar{v} = (\max_{b \subseteq \{1,\ldots,m\}} v_i(b))_{i \in S} \in \mathbb{R}^n$, by McDiarmid's inequality.

Combined with Theorem 2, we get our main guarantee for the class of bundling-boosted auctions. For readability, we state our guarantees assuming $\Delta_i(S_{par}) = 0$ for every i.

Theorem 8. *Let $\mathcal{M}$ be the class of bundling-boosted auctions. Let $N \geq N(\varepsilon, \delta, \mathrm{Pdim}(\mathcal{M}))$ and run LWI with parameters N, p, q. As long as $W(S)^2 \geq ||\bar{v}||_2^2 \ln(1/\delta)/2\eta^2 p^2$,* $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W(S)((1-\eta)p - \varepsilon) - 2\tau_{\mathcal{M}}(q, S_{par})$ *with probability* $\geq 1 - 3\delta$ *conditional on* $\Delta_i(S_{par}) = 0$ *for all* i.

Removing the assumption on $\Delta_i(S_{par})$ would replace the $(1-\eta)p$ loss term with $(1-\eta)(p - \sum_i \Delta_i(S_{par})/W(S))$.

4 Efficient Learning Within an Instance

We now explore two mechanism classes for which LWI can be implemented efficiently by leveraging efficient routines for solving winner determination (a generalization of the problem of computing efficient allocations). Though computing $\widehat{M} = \mathrm{argmax}_{M \in \mathcal{M}} \frac{1}{N} \sum_{t=1}^{N} \mathsf{Rev}_M(S_t)$ is NP-hard since it involves solving winner determination (which is well known to be NP-hard) winner determination can be solved efficiently in practice [Sandholm *et al.*, 2005].

For the class of bundling-VCG auctions, we show that the branch-and-bound technique of Kroer and Sandholm [2015] is compatible with LWI. We did not derive a revenue-guarantee for this class of auctions, however. For the class of *sparse bundling-boosted auctions*, which are bundling-boosted auctions with a constant number of positive boosts, we show that a revenue guarantee similar to (but more sample efficient than) Theorem 8 holds. We then show how LWI can be efficiently implemented for this class.

4.1 Bundling-VCG Auctions

Kroer and Sandholm [2015] give a branch-and-bound algorithm to compute the revenue-maximizing bundling-VCG auction for a given set of bidders. While our setting is different than theirs, their integer-program techniques can be directly used by LWI. Let f denote a function used as an upper bound in branch-and-bound to compute the optimal bundling. For learning groups $S_1, \ldots, S_N$ and x a node in the search tree (corresponding to a partial bundling), let $\hat{f}(x) = \frac{1}{N} \sum_t f(x; S_t)$. Recall that f is *admissible* if its value at any node is an upper bound for the maximum revenue obtainable in the subtree rooted at that node, and f is *monotonic* if it decreases down each path in the search tree. These properties ensure that branch-and-bound finds the revenue-optimal bundling.

Proposition 2. *If f is admissible for computing the optimal bundling, $\hat{f}$ is admissible for computing the empirically optimal bundling. The same holds for monotonicity.*

4.2 Sparse Bundling-Boosted Auctions

Let $\Phi_0 \subset \Phi$ with $|\Phi_0| = B$, and let m_0 be the number of bundles in the finest bundling in Φ_0. Consider the subclass of bundling-boosted auctions for which $\omega(\phi) > 0$ only if $\phi \in \Phi_0$ (and $\omega(\phi) = 0$ otherwise), which we call Φ_0-bundling-boosted auctions. The same argument used to prove Theorem 5 shows that the class of Φ_0-bundling-boosted auctions is $(B, B^2 + m_0 B^3)$-delineable. Let $W^{\Phi_0}(S)$ denote the welfare of the welfare-maximizing allocation to bidders in S, subject to the constraint that the finest bundling respected by the allocation is in Φ_0. The same arguments used to obtain Theorem 7 yield a guarantee with respect to $W^{\Phi_0}(S)$.

Theorem 9. *Let $\mathcal{M}$ be the class of Φ_0-bundling-boosted auctions. Let $N \geq N(\varepsilon, \delta, \mathrm{Pdim}(\mathcal{M}))$ and run LWI with parameters N, p, q. As long as $W^{\Phi_0}(S)^2 \geq ||\bar{v}||_2^2 \ln(1/\delta)/2\eta^2 p^2$, $\mathsf{Rev}_{\widehat{M}}(S_{par}) \geq W^{\Phi_0}(S)((1-\eta)p - \varepsilon) - 2\tau_{\mathcal{M}}(q, S_{par})$ with probability $\geq 1 - 3\delta$ conditional on $\Delta_i(S_{par}) = 0$ for all i.*

For B a fixed constant, the number of learning groups N required by LWI is $O(B \ln(m_0 B))$ (hiding the dependence on ε and δ). In contrast, optimizing over the entire class of bundling-boosted auctions as in Theorem 8 would require N to be exponential (in m). For this class of auctions, we describe an algorithm that implements LWI with run-time exponential in B but polynomial in all other parameters (including the run time of the winner determination routine used). A similar algorithm was used in Balcan et al. [2020], though in a different setting than ours.

Theorem 10. *Let $B = |\Phi_0|$, and let m_0 be the number of bundles in the finest bundling in Φ_0. Given learning groups $S_1, \ldots, S_N$, the empirical-revenue maximizing Φ_0-bundling-boosted auction can be computed in $(Nm_0B)^{O(B)} + 2w(m_0, n)Nm_0B$ time, where $w(m_0, n)$ is the time required to solve winner determination for n buyers with valuations over m_0 items.*

Proof sketch. We first show that there is a set $\mathcal{H}$ of at most $NB^2 + Nm_0B^2$ hyperplanes partitioning $\mathbb{R}^B$ such that *empirical* revenue is linear in ω within each region. Writing these hyperplanes down requires at most $NB + Nm_0B$ calls to our winner determination routine. The maximum empirical revenue in each region can be found by solving a linear program with B variables and at most $|\mathcal{H}|$ constraints. □

4.3 Structural Revenue Maximization

Suppose the mechanism designer can only sample a limited number N of learning groups (due to a run-time constrant, for example). We introduced several new auction classes, but which one should the mechanism designer use in conjunction with LWI? *Structural revenue maximization* (SRM) helps answer this question. SRM suggests maximizing empirical revenue minus a regularization term that penalizes more complex mechanisms to ensure that the chosen auction is indeed likely to generalize well, rather than overfitting to the learning groups. Our generalization guarantee in Theorem 1 provides the appropriate regularizer $\varepsilon_{\mathcal{M}}(N, \delta) = 240\sqrt{\mathrm{Pdim}(\mathcal{M})/N} + \sqrt{2\ln(1/\delta)/N}$. Say the mechanism designer is deciding between auctions in $\mathcal{M}_1$ and auctions in $\mathcal{M}_2$. Let $\widehat{M}_1$, $\widehat{M}_2$ be the empirical-revenue-maximizing auctions from $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, for one run of LWI. The mechanism designer should use mechanism $\widehat{M}_k$, $k \in \{1, 2\}$, that maximizes $\frac{1}{N}\sum_t \mathsf{Rev}_{\widehat{M}_k}(S_t) - \varepsilon_{\mathcal{M}_k}(N, \delta)$ since empirical revenue minus $\varepsilon_{\mathcal{M}}(N, \delta)$ is a more accurate lower bound on expected revenue than empirical revenue alone. An SRM approach combined with LWI is incentive compatible since the final mechanism only depends on the learning groups of bidders. Our use of SRM is similar to SRM across instances, which was discussed in Balcan et al. [2018]. SRM for auction design was first proposed by Balcan et al. [2005], also for learning within an instance (but for unlimited supply).

5 Conclusions and Future Research

We developed a new framework for designing truthful, high-revenue combinatorial auctions for limited supply. Our mechanism *learns within an instance*. It generalizes and improves over previously-studied random-sampling mechanisms.

We proved guarantees on the performance of LWI based on a market-shrinkage term and a new complexity measure we coined *partition discrepancy*, which simultaneously measured intrinsic complexity of the auction class and uniformity in the set of bidders. We explored examples and proved a general bound on partition discrepancy. We then introduced new population-size-independent auction classes, and proved strong generalization bounds for these classes. We showed how LWI can be implemented efficiently by leveraging practically-efficient routines for solving winner determination, and showed how structural revenue maximization helps choose the right auction class to prevent overfitting.

Many interesting new directions arise from this work. First, developing a fuller picture of partition discrepancy would be of independent interest both in mechanism design and learning theory more broadly. Next, population-size-independent auctions are more versatile than other auction formats since they may be designed and specified without knowledge of the number of bidders participating. A further study of such auctions could yield insights into other revenue-maximization settings (for example, when the distribution over bidders is known) when the number of bidders is unknown. Finally, LWI-like frameworks could be applicable to more general mechanism design settings that involve optimization subject to incentive-compatibility constraints.

Acknowledgements

This material is based on work supported by the NSF under grants IIS-1618714, IIS-1718457, IIS-1901403, and CCF-1733556, CCF-1535967, CCF-1910321, and SES-1919453, the ARO under award W911NF2010081, DARPA under cooperative agreement HR00112020003, an AWS Machine Learning Research Award, an Amazon Research Award, a Bloomberg Research Grant, and a Microsoft Research Faculty Fellowship.

References

[Alaei *et al.*, 2009] Saeed Alaei, Azarakhsh Malekian, and Aravind Srinivasan. On random sampling auctions for digital goods. In *EC*, 2009.

[Anthony and Bartlett, 2009] Martin Anthony and Peter L Bartlett. *Neural Network Learning: Theoretical Foundations*. Cambridge University Press, 2009.

[Aumann, 1964] Robert J Aumann. Markets with a continuum of traders. *Econometrica*, 1964.

[Balcan *et al.*, 2005] Maria-Florina Balcan, Avrim Blum, Jason D Hartline, and Yishay Mansour. Mechanism design via machine learning. In *FOCS*, 2005.

[Balcan *et al.*, 2007] Maria-Florina Balcan, Nikhil Devanur, Jason D Hartline, and Kunal Talwar. Random sampling auctions for limited supply. 2007. Technical report, Carnegie Mellon University.

[Balcan *et al.*, 2008] Maria-Florina Balcan, Avrim Blum, and Yishay Mansour. Item pricing for revenue maximization. In *EC*, 2008.

[Balcan *et al.*, 2018] Maria-Florina Balcan, Tuomas Sandholm, and Ellen Vitercik. A general theory of sample complexity for multi-item profit maximization. In *EC*, 2018.

[Balcan *et al.*, 2020] Maria-Florina Balcan, Siddharth Prasad, and Tuomas Sandholm. Efficient algorithms for learning revenue-maximizing two-part tariffs. In *IJCAI*, 2020.

[Barman and Echenique, 2020] Siddharth Barman and Federico Echenique. The edgeworth conjecture with small coalitions and approximate equilibria in large economies. In *EC*, 2020.

[Berend and Tassa, 2010] Daniel Berend and Tamir Tassa. Improved bounds on Bell numbers and on moments of sums of random variables. *Probability and Mathematical Statistics*, 2010.

[Chakraborty *et al.*, 2013] Tanmoy Chakraborty, Zhiyi Huang, and Sanjeev Khanna. Dynamic and nonuniform pricing strategies for revenue maximization. *SIAM Journal on Computing*, 2013.

[Clarke, 1971] Ed H. Clarke. Multipart pricing of public goods. *Public Choice*, 1971.

[Conitzer and Sandholm, 2002] Vincent Conitzer and Tuomas Sandholm. Complexity of mechanism design. In *UAI*, 2002.

[Cramton *et al.*, 2006] Peter Cramton, Yoav Shoham, and Richard Steinberg. *Combinatorial Auctions*. MIT Press, 2006.

[Debreu and Scarf, 1963] Gerard Debreu and Herbert Scarf. A limit theorem on the core of an economy. *International Economic Review*, 1963.

[Devanur and Hartline, 2009] Nikhil R Devanur and Jason D Hartline. Limited and online supply and the Bayesian foundations of prior-free mechanism design. In *EC*, 2009.

[Devanur *et al.*, 2015] Nikhil R Devanur, Jason D Hartline, and Qiqi Yan. Envy freedom and prior-free mechanism design. *Journal of Economic Theory*, 2015.

[Dudley, 1987] Richard M Dudley. Universal Donsker classes and metric entropy. *The Annals of Probability*, 1987.

[Goldberg *et al.*, 2001] Andrew V Goldberg, Jason D Hartline, and Andrew Wright. Competitive auctions and digital goods. In *SODA*, 2001.

[Groves, 1973] Theodore Groves. Incentives in teams. *Econometrica*, 1973.

[Haussler, 1995] David Haussler. Sphere packing numbers for subsets of the boolean n-cube with bounded Vapnik-Chervonenkis dimension. *J. Comb. Theory, Ser. A*, 1995.

[Jehiel *et al.*, 2007] Philippe Jehiel, Moritz Meyer-Ter-Vehn, and Benny Moldovanu. Mixed bundling auctions. *Journal of Economic Theory*, 2007.

[Kroer and Sandholm, 2015] Christian Kroer and Tuomas Sandholm. Computational bundling for auctions. In *AAMAS*, 2015.

[Likhodedov and Sandholm, 2004] Anton Likhodedov and Tuomas Sandholm. Methods for boosting revenue in combinatorial auctions. In *AAAI*, 2004.

[Likhodedov and Sandholm, 2005] Anton Likhodedov and Tuomas Sandholm. Approximating revenue-maximizing combinatorial auctions. In *AAAI*, 2005.

[Mohri and Medina, 2014] Mehryar Mohri and Andres Muñoz Medina. Learning theory and algorithms for revenue optimization in second price auctions with reserve. In *ICML*, 2014.

[Morgenstern and Roughgarden, 2015] Jamie H Morgenstern and Tim Roughgarden. On the pseudo-dimension of nearly optimal auctions. In *NIPS*, 2015.

[Roberts, 1979] Kevin Roberts. The characterization of implementable social choice rules. In J-J Laffont, editor, *Aggregation and Revelation of Preferences*. 1979.

[Sandholm and Likhodedov, 2015] Tuomas Sandholm and Anton Likhodedov. Automated design of revenue-maximizing combinatorial auctions. *Operations Research*, 2015.

[Sandholm *et al.*, 2005] Tuomas Sandholm, Subhash Suri, Andrew Gilpin, and David Levine. CABOB: A fast optimal algorithm for winner determination in combinatorial auctions. *Management Science*, 2005.

[Sandholm, 2003] Tuomas Sandholm. Automated mechanism design: A new application area for search algorithms. In *CP*, 2003.

[Tang and Sandholm, 2012] Pingzhong Tang and Tuomas Sandholm. Mixed-bundling auctions with reserve prices. In *AAMAS*, 2012.

[Vickrey, 1961] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 1961.

A Omitted results and proofs from Section 2

Empirical Rademacher complexity is defined as

$$R_{\mathcal{M}}(S_1,\ldots,S_N)=\mathop{\mathbf{E}}_{\sigma}\left[\sup_{M\in\mathcal{M}}\frac{1}{N}\sum_{t=1}^{N}\sigma_t\mathsf{Rev}_M(S_t)\right],$$

where σ is chosen uniformly at random from $\{-1,1\}^N$. Expected Rademacher complexity is defined as

$$R_{\mathcal{M}}(N;S_{lrn})=\mathop{\mathbf{E}}_{S_1,\ldots,S_N\sim S_{lrn}}[R_{\mathcal{M}}(S_1,\ldots,S_N)].$$

Our LWI mechanism satisfies a standard uniform convergence guarantee since each learning group S_t is sampled independently and identically from $S_{lrn}:=S\setminus S_{par}$ [Anthony and Bartlett, 2009].

Theorem 11. *Let S_{lrn} denote the learning pool of bidders chosen by a run of LWI. With probability at least $1-\delta$ over the draw of $S_1,\ldots,S_N\sim S_{lrn}$, every mechanism $M\in\mathcal{M}$ satisfies $\mathbf{E}_{S_0\sim S_{lrn}}[\mathsf{Rev}_M(S_0)]\le\frac{1}{N}\sum_{t=1}^{N}\mathsf{Rev}_M(S_t)+2R_{\mathcal{M}}(N;S_{lrn})+W(S)\sqrt{\frac{\ln(1/\delta)}{2N}}$.*

The *pseudodimension* of a mechanism class, denoted by $\mathrm{Pdim}(\mathcal{M})$, is the largest positive integer N for which there exist N inputs $S_1,\ldots,S_N$ and N thresholds $r_1,\ldots,r_N$ such that for every $(a_1,\ldots,a_N)\in\{0,1\}^N$, there exists $M\in\mathcal{M}$ such that $\mathsf{Rev}_M(S_t)\ge r_t$ if and only if $a_t=1$.

Rademacher complexity and pseudodimension are related via the following bound due to Dudley [1987]:

$$R_{\mathcal{M}}(N;S_{lrn})\le 60\cdot W(S)\sqrt{\frac{\mathrm{Pdim}(\mathcal{M})}{N}}.$$

Delineability and pseudodimension are related via the main result of Balcan et al. [2018]: if $\mathcal{M}$ is (d,h)-delineable, $\mathrm{Pdim}(\mathcal{M})\le 9d\ln(4dh)$.

Proof of Theorem 1. Let

$$\varepsilon=\varepsilon_{\mathcal{M}}(\delta,N)=2R_{\mathcal{M}}(N;S_{lrn})+\sqrt{\frac{\ln(1/\delta)}{2N}}.$$

By Theorem 11 it holds with probability at least $1-\delta$ that for all $M\in\mathcal{M}$,

$$\mathop{\mathbf{E}}_{S_0\sim S_{lrn}}[\mathsf{Rev}_M(S_0)]\le\frac{1}{N}\sum_{t=1}^{N}\mathsf{Rev}_M(S_t)+\varepsilon W(S),$$

and symmetrically it holds with probability at least $1-\delta$ that for all $M\in\mathcal{M}$,

$$\frac{1}{N}\sum_{t=1}^{N}\mathsf{Rev}_M(S_t)\le\mathop{\mathbf{E}}_{S_0\sim S_{lrn}}[\mathsf{Rev}_M(S_0)]+\varepsilon W(S).$$

Hence, the probability of both events is at least $1-2\delta$.

Let $M^*=\operatorname{argmax}_{M\in\mathcal{M}}\mathsf{Rev}_M(S_{par})$. For brevity, let $\tau=\tau_{\mathcal{M}}(q,S_{par})$. Then,

$$\begin{aligned}\mathsf{Rev}_{\widehat{M}}(S_{par})&\ge\mathop{\mathbf{E}}_{S_0\sim S_{lrn}}[\mathsf{Rev}_{\widehat{M}}(S_0)]-\tau\\&\ge\frac{1}{N}\sum_{t=1}^{N}\mathsf{Rev}_{\widehat{M}}(S_t)-\varepsilon W(S)-\tau\\&\ge\frac{1}{N}\sum_{t=1}^{N}\mathsf{Rev}_{M^*}(S_t)-\varepsilon W(S)-\tau\\&\ge\mathop{\mathbf{E}}_{S_0\sim S_{lrn}}[\mathsf{Rev}_{M^*}(S_{par})]-2\varepsilon W(S)-\tau\\&\ge\mathsf{Rev}_{M^*}(S_{par})-2\varepsilon W(S)-2\tau\\&=W(S)(L_{\mathcal{M}}(p,S_{par})-2\varepsilon)-2\tau,\end{aligned}$$

as desired. Part (b) is a consequence of the bound $R_{\mathcal{M}}(N;S_{lrn})\le 60W(S)\sqrt{\frac{\mathrm{Pdim}(\mathcal{M})}{N}}$ [Dudley, 1987]. □

Suppose $\mathcal{M}$ is (d,h)-delineable where $h:\mathbb{N}\to\mathbb{N}$ is a convex function of the number of bidders n. That is, for every set S of bidder valuations, there are at most $h(n)$ hyperplanes partitioning $\mathbb{R}^d$ such that $\mathsf{Rev}_S(\theta):=\mathsf{Rev}_\theta(S)$ is linear in θ over each connected component over $\mathbb{R}^d$ determined by the hyperplanes. In this case, the bounds due to Dudley and Balcan et al. do not apply, since h ought to be independent of the number of bidders in order to apply the results of Balcan, Sandholm, and Vitercik [2018] (this is because the revenue functions induced by $\mathcal{M}$ accept any number of bidder valuations as input). We can nevertheless give a (slightly weaker) bound on Rademacher complexity by more directly exploiting the geometric nature of delineability. The assumption that h is convex is simply because that is the case for all the mechanism classes we consider [Balcan *et al.*, 2018]. We will also assume $h(0)=1$ since when there are no bidders the only allocation is for the seller to keep all the goods.

If $\mathcal{M}$ is an auction class that can be parameterized in a population size-independent manner, then note that $\mathcal{M}$ must necessarily satisfy a population size-independent version of delineability as in the main body of the paper. Our analysis here is in the case that one is not able to prove these stronger delineability guarantees. We show that it is still possible to prove guarantees on LWI by directly exploiting delineability to bound the Rademacher complexity.

Lemma 1. *Let $\mathcal{M}$ be (d,h)-delineable, where h is convex in the number of bidders and $h(0)=1$. LWI satisfies*

$$R_{\mathcal{M}}(N;S_{lrn})\le W(S)\sqrt{\frac{4d\ln(Nh(q|S_{lrn}|))}{N}}.$$

Proof. Let $\mathcal{M}$ be parametrized by vectors $\theta\in\mathbb{R}^d$. We have

$$\begin{aligned}R_{\mathcal{M}}(S_1,\ldots,S_N)&=\frac{1}{N}\mathop{\mathbf{E}}_{\sigma}\Big[\sup_{\theta}\sum_{t=1}^{N}\sigma_t\mathsf{Rev}_\theta(S_t)\Big]\\&=\frac{1}{N}\mathop{\mathbf{E}}_{\sigma}\Big[\sup_{\theta}\sum_{t=1}^{N}\sigma_t\mathsf{Rev}_{S_t}(\theta)\Big].\end{aligned}$$

Fix $\sigma_1,\ldots,\sigma_N\in\{-1,1\}$. Let $n_t=|S_t|$. Then, there exist $h(n_1)+\cdots+h(n_N)$ hyperplanes in $\mathbb{R}^d$ such that over each region determined by the hyperplanes, the quantity $\sum_{t=1}^{n} \sigma_t \mathsf{Rev}_{S_t}(\theta)$ is linear in θ. Hence the maximum is achieved at a vertex of one of these regions. The number of such regions is at most $(h(n_1) + \cdots + h(n_N))^d$. Each region is an intersection of at most $(h(n_1) + \cdots + h(n_N))$ halfspaces, and is thus a convex polytope with at most $(h(n_1) + \cdots + h(n_N))^d$ vertices. Let V denote the collection of all the vertices of every region, so

$$|V| \le (h(n_1) + \cdots + h(n_N))^{2d}.$$

For every $\sigma \in \{-1, 1\}^N$, $\sum_{t=1}^{N} \sigma_t \mathsf{Rev}_{S_t}(\theta)$ is maximized at a point in V. Therefore, the supremum over $\theta \in \mathbb{R}^d$ can be replaced with a maximum over $\theta \in V$. Applying Massart's finite class lemma yields

$$\begin{aligned}
&\frac{1}{N} \mathop{\mathbf{E}}_{\sigma} \left[\max_{\theta \in V} \sum_{t=1}^{N} \sigma_t \mathsf{Rev}_{S_t}(\theta) \right] \\
&\quad \le \frac{\max_\theta \sqrt{\mathsf{Rev}_{S_1}(\theta)^2 + \cdots + \mathsf{Rev}_{S_N}(\theta)^2} \cdot \sqrt{2 \ln |V|}}{N} \\
&\quad \le W(S) \cdot \sqrt{\frac{4d \ln(h(n_1) + \cdots + h(n_N))}{N}}.
\end{aligned}$$

Now, let h^{lin} be the linear function with $h^{lin}(0) = 1$ and $h^{lin}(q|S_{lrn}|) = h(q|S_{lrn}|)$. h is convex, so $h(x) \le h^{lin}(x)$ for all $0 \le x \le q|S_{lrn}|$. Then,

$$\begin{aligned}
&\mathop{\mathbf{E}}_{S_1, \ldots, S_N} \left[\sqrt{\frac{4d \ln(h(n_1) + \cdots + h(n_N))}{N}} \right] \\
&\quad \le \sqrt{\frac{4d \ln(N\, \mathbf{E}[h(n_1)])}{N}} \\
&\quad \le \sqrt{\frac{4d \ln(N\, \mathbf{E}[h^{lin}(n_1)])}{N}} \\
&\quad \le \sqrt{\frac{4d \ln(N h(q|S_{lrn}|))}{N}},
\end{aligned}$$

where the first inequality is due to Jensen's inequality and the rest follow from linearity and that $\mathbf{E}[n_1] = q|S_{lrn}|$. $\square$

In case $\mathcal{M}$ satisfies (d, h) delineability, where $h : \mathbb{N} \to \mathbb{N}$ is a convex function of the number of bidders, we get a guarantee using Lemma 1 in place of Dudley's bound [Dudley, 1987] and the main result of Balcan et al. [2018].

Theorem 12. *Let $\mathcal{M}$ be a mechanism class that is (d, h)-delineable, where h is convex in the number of bidders. Let S_{par} denote the participatory set of bidders chosen by a run of LWI. Then, with probability $\ge 1 - 2\delta$ over the draw of $S_1, \ldots, S_N \sim_q S \setminus S_{par}$, $\mathsf{Rev}_{\widehat{M}}(S_{par}) \ge W(S)\Big(L_{\mathcal{M}}(S_{par} - 8\sqrt{\frac{d \ln(N h(q|S_{par}|))}{N}} - \sqrt{\frac{2\ln(1/\delta)}{N}}\Big) - 2\tau_{\mathcal{M}}(p, q, S_{par})$.*

Partition discrepancy for replica economies

Suppose $S_0 = \{v_1, v_2, v_3\}$, and S consists of n_0 replicas of S_0. Let $\mathcal{M}$ be any auction class that can be parameterized in a way that does not depend on the number of bidders (for example, any of the auctions we define in Section 3). Any $M \in \mathcal{M}$ can be identified by the overall allocation it uses, and the at most m allocations it uses to determine payments. Hence, M can be encoded as a vector of length at most $(m+1)^2$ (there are at most $m+1$ allocations, and each sells to at most m buyers). Over all $M \in \mathcal{M}$, the number of such vectors is at most $3^{(m+1)^2}$, since bidders of the same type are indistinguishable. Suppose LWI is run with $p = 1/2, q = 1$. Due to Chernoff bounds, for n_0 large, it holds with exceedingly high probability that the number of bidders of each type (v_1, v_2, or v_3) in S_{par} is in $[n_0/2 - 10\sqrt{n_0}, n_0/2 + 10\sqrt{n_0}]$ (and likewise for $S \setminus S_{par}$). So if n_0 is sufficiently large, both S_{par} and $S \setminus S_{par}$ contain enough bidders of each type to form each of the at most $3^{(m+1)^2}$ auction vectors with extremely high probability, in which case $\mathsf{Rev}_M(S_{par}) = \mathsf{Rev}_M(S \setminus S_{par})$ for every $M \in \mathcal{M}$ and so $\tau_{\mathcal{M}}(1, S_{par}) = 0$.

General bound on partition discrepancy

Proof of Theorem 3. Let $S = \{v_1, \ldots, v_n\}$ and let $Y_i = 1$ with probability $1/3$ and $Y_i = 0$ with probability $2/3$. Fix a single mechanism $M \in \mathcal{M}$. Let $g(Y_1, \ldots, Y_n) = \mathsf{Rev}_M(S_{par}) - \mathbf{E}_{S_0 \sim S_{lrn}}[\mathsf{Rev}_M(S_0)]$ where $S_{par} = \{v_i : Y_i = 1\}$. Let c_i be an upper bound for $|g(Y_1, \ldots, Y_n) - g(Y_1, \ldots, Y_i', \ldots, Y_n)|$. We have $\sum_i c_i^2 \le ||\tilde{v}||_2^2$. By symmetry and the fact that each bidder has an equal probability of $1/3$ of being in S_{par} or in S_0, $\mathbf{E}_{S_{par} \sim S, S_0 \sim S_{lrn}}[\mathsf{Rev}_M(S_{par}) - \mathsf{Rev}_M(S_0)] = 0$. McDiarmid's inequality therefore yields,

$$\begin{aligned}
&\Pr_{S_{par} \sim S}(\tau_M(1/2, S_{par}) \ge t) \\
&\quad = \Pr_{S_{par} \sim S}(|\mathsf{Rev}_M(S_{par}) - \mathop{\mathbf{E}}_{S_0 \sim S_{lrn}}[\mathsf{Rev}_M(S_0)]| \ge t) \\
&\quad \le e^{-2t^2/\sum c_i^2}.
\end{aligned}$$

Fix t. Let $\widetilde{\mathcal{M}} \subseteq \mathcal{M}$ be a subset of mechanisms that forms a $t/2^{n+2}$ L_1-cover of $\{(\mathsf{Rev}_M(S'))_{S' \subseteq S} : M \in \mathcal{M}\} \subset \mathbb{R}^{2^n}$. Now, for each $M \in \mathcal{M}$, there is $\widetilde{M} \in \widetilde{\mathcal{M}}$ such that for any S_{par}, with $n_1 = |S_{par}|$, $n_2 = |S_{lrn}|$, we have

$$\begin{aligned}
&\tau_M(1/2, S_{par}) \\
&\quad = \left| \mathsf{Rev}_M(S_{par}) - \frac{1}{2^{n_2}} \sum_{S_0 \subseteq S_{lrn}} \mathsf{Rev}_M(S_0) \right| \\
&\quad \le |\mathsf{Rev}_M(S_{par}) - \mathsf{Rev}_{\widetilde{M}}(S_{par})| \\
&\qquad + \frac{1}{2^{n_2}} \sum_{S_0 \subseteq S_{lrn}} |\mathsf{Rev}_M(S_0) - \mathsf{Rev}_{\widetilde{M}}(S_0)| \\
&\qquad + \left| \mathsf{Rev}_{\widetilde{M}}(S_{par}) - \frac{1}{2^{n_2}} \sum_{S_0 \subseteq S_{lrn}} \mathsf{Rev}_{\widetilde{M}}(S_0) \right| \\
&\quad \le \frac{t}{4} + \frac{t}{2^{n_2+2}} + \tau_{\widetilde{M}}(1/2, S_{par}) \\
&\quad \le \frac{t}{2} + \tau_{\widetilde{M}}(1/2, S_{par}).
\end{aligned}$$

A union bound yields

$$\begin{aligned}
&\Pr\Big(\sup_{M\in\mathcal{M}}\tau_M(1/2,S_{par})\geq t\Big)\\
&\quad\leq \Pr\Big(\max_{\widetilde{M}\in\widetilde{\mathcal{M}}}\tau_{\widetilde{M}}(1/2,S_{par})\geq t/2\Big)\\
&\quad\leq \mathcal{N}_1(t/2^{n+2};\mathcal{M};2^n)\cdot e^{-\frac{t^2}{2\sum c_i^2}}\\
&\quad\leq e(\mathrm{Pdim}(\mathcal{M})+1)\left(\frac{2^{n+3}eW(S)}{t}\right)^{\mathrm{Pdim}(\mathcal{M})}\cdot e^{-\frac{t^2}{2\sum c_i^2}}.
\end{aligned}$$

The final inequality follows from a well-known L_1-covering-number bound in terms of pseudodimension [Haussler, 1995; Anthony and Bartlett, 2009]. Taking t at least $||\tilde{v}||_2\sqrt{2n\,\mathrm{Pdim}(\mathcal{M})\ln\frac{4e^2\,\mathrm{Pdim}(\mathcal{M})W(S)}{\delta}}$ yields the desired confidence of at least $1-\delta$. $\square$

B Omitted results and proofs from Section 3

Proof of Proposition 1. Consider running a second-price auction on the grand bundle with reserve price $v_{\alpha n}$. If any bidder who values the grand bundle at least $v_{\alpha n}$ is in S_{par}, the revenue obtained is at least $v_{\alpha n}$. This event occurs with probability $\Pr(\cup_{i\leq\alpha n}\{i\in S_{par}\})\geq 1-(1-p)^{\alpha n}\geq 1-e^{-\alpha np}$. $\square$

Recall that for a bundling ϕ, ϕ-VCG allocations refer to VCG allocations subject to the constraint that bundles in ϕ are treated as indivisible items.

Proof of Theorem 5. We prove the result for bundling-boosted auctions. Fix the input set of bidders. For each bundling ϕ, let $\beta^\phi=\operatorname{argmax}_{\beta:\mathsf{blg}(\beta)=\phi}W(\beta)$, and let $m(\phi)$ denote the number of bundles in ϕ. Note that if β^ϕ is not the ϕ-VCG allocation, then for ϕ' the coarsest bundling respected by the ϕ-VCG allocation, $\beta^{\phi'}$ is the VCG allocation corresponding to both ϕ and ϕ' (ϕ' can be obtained by combining certain bundles in ϕ).

Now, for any bundling-boosted auction parameters ω, the overall allocation chosen must be one of $\{\beta^\phi:\phi\in\Phi\}$ (which contains as a subset the collection of all bundling VCG allocations, as remarked above). This is because for any allocation α, α is given the same boost as $\beta^{\mathsf{blg}(\alpha)}$, which has greater welfare by definition.

We now count the total number allocations that can ever be used by ω when any bidder is absent. For a given $\phi\in\Phi$, exactly $m(\phi)$ bidders are allocated any items by β^ϕ. If bidder i is not allocated any items by β^ϕ, then β^ϕ also maximizes welfare among all allocations α such that $\mathsf{blg}(\alpha)=\phi$ when bidder i is absent. If bidder i is allocated something by β^ϕ, let $\beta^\phi_{-i}=\operatorname{argmax}_{\beta:\mathsf{blg}(\beta)=\phi}W_{-i}(\beta)$. For any setting of the parameters ω, the allocation used when bidder i is absent will be of the form β^ϕ_{-i}. This is because for any allocation α,

$$W_{-i}(\alpha)+\omega(\mathsf{blg}(\alpha))\leq W_{-i}\big(\beta_{-i}^{\mathsf{blg}(\alpha)}\big)+\omega(\mathsf{blg}(\alpha)).$$

There are a total of $m(\phi)$ such unique allocations (not including β^ϕ). Now, if bidder i is not allocated any items by β^ϕ for any $\phi\in\Phi$, the absence of bidder i does not change the allocation used by any $\omega\in\mathbb{R}^{|\Phi|}$. The total number of bidders whose absence can change the allocation used is thus at most

$$\sum_{\phi\in\Phi}m(\phi)\leq m|\Phi|.$$

Finally, we count the number of hyperplanes partitioning the parameter space such that the allocations used are constant within any region. There are $\binom{|\Phi|}{2}$ hyperplanes of the form

$$W(\beta^\phi)+\omega(\phi)=W(\beta^{\phi'})+\omega(\phi')$$

for every $\phi,\phi'\in\Phi$. For each of the at most $m|\Phi|$ bidders whose absence potentially changes the allocation used, there are at most $\binom{|\Phi|}{2}$ hyperplanes of the form

$$W(\beta^\phi_{-i})+\omega(\phi)=W(\beta^{\phi'}_{-i})+\omega(\phi')$$

for every $\phi,\phi'\in\Phi$ such that i is allocated a nonempty bundle by β^ϕ and $\beta^{\phi'}$. The total number of hyperplanes is thus at most

$$\binom{|\Phi|}{2}+m|\Phi|\binom{|\Phi|}{2}<|\Phi|^2+m|\Phi|^3.$$

$\square$

Proof of Theorem 6. For a given bundle-size-boosted auction, consider the equivalent bundling-boosted auction. Any two bundlings that are indistinguishable with respect to the sizes of their bundles are given the same boost, since $\omega(\phi)=\sum_{b\in\phi}\omega(|b|)$. Hence, we may run the same argument of Theorem 5 but instead of considering all $|\Phi|$ bundlings we only need to consider the number $p(m)$ of integer partitions of m. It is well known that there is a constant B such that $p(m)<e^{B\sqrt{m}}$. $\square$

Proof of Theorem 7. Let $\phi=\mathsf{blg}(\beta)$ and for each i let $\phi_{-i}=\mathsf{blg}(\beta_{-i})$. Consider the bundling-boosted auction with $\omega(\phi)=0$, $\omega(\phi_{-i})=W(\beta)-W(\beta_{-i})$, and $\omega(\phi')=-\infty$ for all other bundlings $\phi'\in\Phi\setminus\{\phi,\phi_{-1},\dots,\phi_{-n}\}$.

We show that when $\Delta_i(S)=0$ for all i, this auction extracts revenue equal to $W(S)$. The proof of the more general statement in terms of $\Delta_i(S)$ is similar; we describe the necessary modifications at the end. First we show that β is the overall allocation chosen by this auction. For any α such that $\mathsf{blg}(\alpha)\notin\{\phi,\phi_{-1},\dots,\phi_{-n}\}$, $\lambda(\alpha)=-\infty$, so such allocations are never chosen.

Case 1. $\mathsf{blg}(\alpha)=\phi$. Then as $\lambda(\beta)=\lambda(\alpha)=\omega(\phi)=0$, $W(\beta)+\lambda(\beta)\geq W(\alpha)+\lambda(\alpha)$, since β is an efficient allocation.

Case 2. $\mathsf{blg}(\alpha)=\phi_{-i}$ for some i. Then $\lambda(\alpha)=\omega(\phi_{-i})$ and $W(\beta_{-i})\geq W(\alpha)$, so $W(\beta)+\lambda(\beta)=W(\beta)=W(\alpha)+(W(\beta)-W(\alpha))\geq W(\alpha)+(W(\beta)-W(\beta_{-i}))=W(\alpha)+\lambda(\alpha)$ for all i.

Hence β is the overall allocation used. Next we show that when bidder i is absent, β_{-i} is the allocation used by this auction.

Case 1. $\mathsf{blg}(\alpha)=\phi$. Then $W_{-i}(\beta_{-i})+\lambda(\beta_{-i})=W(\beta)\geq W(\alpha)\geq W(\alpha)-v_i(\alpha)=W_{-i}(\alpha)+\lambda(\alpha)$.

Case 2. $\mathsf{blg}(\alpha)=\phi_{-i}$. Then $\lambda(\beta_{-i})=\lambda(\alpha)=\omega(\phi_{-i})$, so $W_{-i}(\beta_{-i})+\lambda(\beta_{-i})\geq W_{-i}(\alpha)+\lambda(\alpha)$.

Case 3. $\mathsf{blg}(\alpha) = \phi_{-k}$ for $k \neq i$. Then $W(\beta_{-k}) \geq W(\alpha)$, so $W_{-i}(\alpha)+\lambda(\alpha) = W_{-i}(\alpha)+W(\beta)-W(\beta_{-k}) \leq W_{-i}(\alpha)+W(\beta) - W(\alpha) = W(\beta) - v_i(\alpha) \leq W(\beta) = W(\beta_{-i}) + \lambda(\beta_{-i})$.

Hence β_{-i} is the allocation used when bidder i is not present. We have shown that the allocations used by this bundling-boosted auction are precisely the VCG allocations. The payment of bidder i is therefore $(W_{-i}(\beta_{-i})+\lambda(\beta_{-i})) - (W_{-i}(\beta) + \lambda(\beta)) = W(\beta) - (W(\beta) - v_i(\beta)) = v_i(\beta)$, and so the total revenue is $W(\beta)$. The proof of the general statement is similar. The only difference is that ω might not use the VCG allocations, but is guaranteed to use allocations for which the boosted welfare does not differ much from that of the corresponding VCG allocation by much. □

We give a simple example of bidder valuations that satisfy $\Delta_i(S') = 0$ for every i and every $S' \subseteq S$ involving bidders whose "most desired" bundles intersect.

Example 4. *Let $b_1, \ldots, b_n \subseteq \{1, \ldots, m\}$ be distinct bundles such that $b_i \cap b_j \neq \emptyset$ for every i, j. Let $c_1 > c_2 > \cdots > c_n > 0$, and let $\varepsilon > 0$ be sufficiently small. The valuation of bidder i satisfies $v_i(b) = c_i$ if $b \supseteq b_i$ and $v_i(b) \leq \varepsilon$ otherwise. Then, for any subset of the bidders $S' \subseteq \{1, \ldots, n\}$, the welfare-maximizing allocation gives bundle b_i to $i = \min(S')$ and allocates the remaining items to the other bidders. When i is absent, the welfare-maximizing allocation gives bundle $b_{i'}$ to $i' = \min(S' \setminus \{i\})$ and allocates the remaining items to the other bidders. This is clearly the welfare-maximizing allocation among all allocations respecting the same finest bundling. Now, for each $j \neq i$, if the finest bundlings respected by each of the welfare-maximizing allocations when j is absent are all distinct, then $\Delta_i(S') = 0$ for every i, S'. (since the welfare-maximizing allocation gives b_i to i and uses a distinct bundling on the remaining buyers for each j). Otherwise, $\Delta_i(S')$ is nevertheless small, since the welfare extracted from bidders excluding i when any bidder $j \neq i$ is absent is small.*

C Omitted proofs from Section 4

Proof of Proposition 2. Suppose we are at node x of a branch-and-bound computation (representing a partial bundling). Let T_x denote the subtree rooted at x. Then, $\hat{f}(x) = \frac{1}{N}\sum_{t=1}^{N} f(x; S_t) \geq \frac{1}{N}\sum_{t=1}^{N} \max_{\phi \in T_x} \mathsf{Rev}_\phi(S_t) \geq \max_{\phi \in T_x} \frac{1}{N}\sum_{t=1}^{N} \mathsf{Rev}_\phi(S_t)$. Monotonicity of $\hat{f}$ is also immediate. □

Proof of Theorem 10. For a bundling $\phi \in \Phi$, let $\mathsf{Winner}(\phi, S_t)$ denote the welfare-maximizing allocation among bidders in S_t respecting bundling ϕ. Let $\beta^{\phi,t} = \mathsf{Winner}(\phi, S_t)$ and let $S_t^\phi \subseteq S_t$ be the set of bidders in S_t who get allocated a nonempty bundle by $\beta^{\phi,t}$. For each bidder $i \in S_t^\phi$, let $\beta_{-i}^{\phi,t} = \mathsf{Winner}(\phi, S_t \setminus \{i\})$. The proof of Theorem 5 shows that any ω must use a subset of these allocations. Winner is called at most $NB + Nm_0B$ times, since $|S_t^\phi| \leq m_0$.

For each t and each pair of bundlings ϕ, ϕ', let $H(t, \phi, \phi')$ denote the hyperplane

$$\sum_{i \in S_t} v_i(\beta^{\phi,t}) + \omega(\phi) = \sum_{i \in S_t} v_i(\beta^{\phi',t}) + \omega(\phi'),$$

and for each $i \in S_t^\phi$ let $H_{-i}(t, \phi, \phi')$ denote the hyperplane

$$\sum_{j \in S_t \setminus \{i\}} v_j(\beta_{-i}^{\phi,t}) + \omega(\phi) = \sum_{j \in S_t \setminus \{i\}} v_j(\beta_{-i}^{\phi',t}) + \omega(\phi').$$

Let $\mathcal{H}$ denote the collection of these hyperplanes. The total number of such hyperplanes is at most $NB^2 + Nm_0B^2$ (as argued in Theorem 5). It is a basic combinatorial fact that $\mathcal{H}$ partitions $\mathbb{R}^B$ into at most $|\mathcal{H}|^B \leq (NB^2 + Nm_0B^2)^B \leq (2Nm_0B^2)^B$ regions (each region is an intersection of at most $|\mathcal{H}|$ halfspaces). The empirical revenue over $S_1, \ldots, S_N$ is linear in ω over each region, since the allocations of the bundling-boosted auction remain constant within each region. Thus, the maximum empirical revenue can be found by solving a linear program within each region. Each linear program involves B variables and at most $|\mathcal{H}|$ constraints and can thus be solved in $poly(|\mathcal{H}|, B)$ time.

A representation of each of the regions as a $0/1$ vector of length $|\mathcal{H}|$ can be computed in $poly(|\mathcal{H}|^B)$ time using the following high-level procedure. Initialize the list of regions as $\mathbb{R}^B$ (represented by the empty set of constraints). Iterate over the set of hyperplanes. For each hyperplane, check whether it intersects any region in the current list of regions. If so, update each region it intersects by adding the two constraints corresponding to the two new halfspaces. Our algorithm solves the corresponding linear program for every such region and picks the solution that yields highest empirical revenue. □