Reducing Mechanism Design to Algorithm Design via Machine Learning ⋆

Maria-Florina Balcan [1]

Carnegie Mellon University, Pittsburgh, PA 15213.

Avrim Blum [1]

Carnegie Mellon University, Pittsburgh, PA 15213.

Jason D. Hartline

Microsoft Research, Mountain View, CA 94043.

Yishay Mansour [2]

School of Computer Science, Tel-Aviv University.

Abstract

We use techniques from sample-complexity in machine learning to reduce problems of incentive-compatible mechanism design to standard algorithmic questions, for a broad class of revenue-maximizing pricing problems. Our reductions imply that for these problems, given an optimal (or β-approximation) algorithm for an algorithmic pricing problem, we can convert it into a $(1+\epsilon)$-approximation (or $\beta(1+\epsilon)$-approximation) for the incentive-compatible mechanism design problem, so long as the number of bidders is sufficiently large as a function of an appropriate measure of complexity of the class of allowable pricings. We apply these results to the problem of auctioning a digital good, to the *attribute auction* problem which includes a wide variety of discriminatory pricing problems, and to the problem of item-pricing in unlimited-supply combinatorial auctions. From a machine learning perspective, these settings present several challenges: in particular, the "loss function" is discontinuous, is asymmetric, and has a large range. We address these issues in part by introducing a new form of covering-number bound that is especially well-suited to these problems and may be of independent interest.

Key words: Mechanism Design, Machine Learning, Sample Complexity, Profit Maximization, Unlimited Supply, Digital Good Auction, Attribute Auctions, Combinatorial Auctions, Structural Risk Minimization, Covering Numbers.

Preprint submitted to Elsevier Science *8 September 2007*

1 Introduction

In recent years there has been substantial work on problems of algorithmic mechanism design. These problems typically take a form similar to classic algorithm design or approximation-algorithm questions, except that the inputs are each given by *selfish agents* who have their own interest in the outcome of the computation. As a result it is desirable that the mechanisms (the algorithms and protocol) be *incentive compatible* — meaning that it is in each agent's best interest to report its true value — so that agents do not try to game the system. This requirement can greatly complicate the design problem.

In this paper we consider the design of mechanisms for one of the most fundamental economic objectives: *profit maximization*. Agents participating in such a mechanism may choose to falsely report their preferences if it might benefit them. What we show, however, is that so long as the number of agents is sufficiently large as a function of a measure of the complexity of the mechanism design problem, we can apply sample-complexity techniques from learning theory to reduce this problem to standard algorithmic questions in a broad class of settings. It is useful to think of the techniques we develop in the context of designing an auction to sell some goods or services, though they also apply in more general scenarios.

In a seminal paper Myerson [33] derives the optimal auction for selling a single item given that the bidders' true valuations for the item come from some known *prior distribution*. His mechanism generalizes trivially to any single-parameter agent setting with arbitrary supply constraints or costs to the auctioneer for the outcome produced. Following a trend in the recent computer science literature on optimal auction design, we consider the *prior-free* setting in which there is no underlying distribution on valuations and we wish to perform well for any (sufficiently large) set of bidders. In absence of a known prior distribution we will use machine learning techniques to estimate properties of the bidders' valuations. We consider the *unlimited supply* setting in which this

⋆ A preliminary version of this paper appears in Proceedings of the 46th Annual Symposium on Foundations of Computer Science (FOCS) 2005, under the title "Mechanism Design via Machine Learning".

Email addresses: ninamf@cs.cmu.edu (Maria-Florina Balcan), avrim@cs.cmu.edu (Avrim Blum), hartline@microsoft.com (Jason D. Hartline), mansour@cs.tau.ac.il (Yishay Mansour).

[1] Research supported in part by NSF grants CCR-0105488, CCR-0122581, ITR IIS-0121678.

[2] This work was supported in part by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778, by a grant no. 1079/04 from the Israel Science Foundation, by a grant from BSF and an IBM faculty award. This publication reflects the authors' views only.

problem is conceptually simpler because there are no infeasible allocations; though, it is often possible to obtain results for limited supply or with cost functions on the outcome via reduction to the unlimited supply case [25,19,2]. Research in optimal prior-free auction design is important for optimal auction design because it directly links inaccurate distributional knowledge typical of small markets with loss in performance.

Implicit in mechanism design problems is the fact that the selfish agents that will be participating in the mechanism have *private information* that is known only to them. Often this private information is simply the agent's valuation over the possible outcomes the mechanism could produce. For example, when selling a single item (with the standard assumption that an agent only cares if they get the item or not and not whether another agent gets it) this valuation is simply how much they are willing to pay for the item. There may also be *public information* associated with each agent. This information is assumed to be available to the mechanism. Such information is present in structured optimization problems such as the *knapsack auction problem* [2] and *multicast auction problem* [19] and is the natural way to generalize optimal auction design for independent but non-identically distributed prior distributions (which are considered by Myerson [33]) to the prior-free setting. There are many standard economic settings where such public information is available, e.g., in the college tuition mechanism, in-state or out-of-state residential status is public; for acquiring a loan, a consumer's credit report is public information; for automobile insurance, driving records, credit reports, and the make and color of the vehicle are public information.

A fundamental building block of an incentive compatible mechanism is an *offer*. For full generality an offer can be viewed as an incentive compatible mechanism for one agent. As an example, if we are selling multiple units of a single item, an offer could be a *take-it-or-leave-it* price per unit. A rational agent would accept such an offer if it is lower than the agent's valuation for the item and reject if it is greater. Notice that if all agents are given the same take-it-or-leave-it price then the outcome is *non-discriminatory* and the same price is paid by all winners. Prior-free auctions based on this type of non-discriminatory pricing have been considered previously (see, e.g., [25]).

One of the main motivations of this work is to explore *discriminatory pricing* in optimal auction design. There are two standard means to achieve discriminatory pricing. The first, is to discriminate based on the public information of the consumer. Naturally, loans are more costly for individuals with poor credit scores, car insurance is more expensive for drivers with points on their driving record, and college tuition at state run universities is cheaper for students that are in-state residents. In this setting a reasonable offer might be a mapping from the public information of the agents to a take-it-or-leave-it price. We refer to these types of offers as *pricing functions*. The second standard means for discriminatory pricing is to introduce similar products of different qualities and price them differently. Consumers who cannot afford the expensive high-quality version may still purchase an inexpensive low-quality version. This practice is common, for example, in software sales, electronics sales, and airline ticket sales. An offer for the multiple good setting could be a take-it-or-leave it price for each good. An agent would then be free to select the good (or bundle of goods) with the (total) price that they most prefer. We refer to these types of offers as *item pricings*.

Notice that allowing offers in the form of pricing functions and item pricings, as described above, provides richness to both algorithmic and mechanism design questions. This richness; however, is not without cost. Our performance bounds are parameterized by a suitable notion of the *complexity* of the class of allowable offers. It is natural that this kind of complexity should affect the ability of a mechanism to optimize. It is easier to approximate the optimal offer from a simple classes of offers, such as take-it-or-leave-it prices for a single item, than it is for a more complex class of offers, such as take-it-or-leave-it prices for multiple items. Our prior-free analysis makes the relationship between a mechanism's performance and the complexity of allowed offers precise.

We phrase our auction problem generically as: given some class of reasonable offers, can we construct an incentive-compatible auction that obtains profit close to the profit obtained by the optimal offer from this class? The auctions we discuss are generalizations of the random sampling auction of Goldberg et al. [26]. These auctions make use of a (non-incentive-compatible) algorithm for computing a best (or approximately best) offer from a given class for any set of consumers. Thus, we can view this construction as reducing the optimal mechanism design problem to the optimal algorithm design problem.

The idea of the reduction is as follows. Let $\mathcal{A}$ be an algorithm (exact or approximate) for the purely algorithmic problem of finding the optimal offer in some class $\mathcal{G}$ for any given set of consumers S with known valuations. Our auction, which does not know the valuations a priori, asks the agents to report their valuations (as bids), splits agents randomly into two sets S_1 and S_2, runs the algorithm $\mathcal{A}$ separately on each set (perhaps adding an additional penalty term to the objective to penalize solutions that are too "complex" according to some measure), and then applies the offer found for S_1 to S_2 and the offer found on S_2 to S_1. The incentive compatibility of this auction allows us to assume that the agents will indeed report their true valuations. Sample-complexity techniques adapted from machine learning theory can then give a guarantee on the quality of the results if the market size is sufficiently large compared to a measure of complexity of the class of possible solutions. From an economics perspective, this can be viewed as replacing the Bayesian assumption that bidders come from a known prior distribution (e.g., as in Myerson's work [33]) with the use of learning, over a random subset S_1 of an arbitrary set of bidders

S, to get enough information to apply to S_2 (and vice versa).

It is easy to see that as the size of the market grows, the law of large numbers indicates that the above approach is asymptotically optimal. This is not surprising as conventional economic wisdom suggests that even the approach of market analysis followed by the Bayesian optimal mechanism would incur negligibly small loss compared to the Bayesian optimal mechanism which was endowed with foreknowledge of the distribution. In contrast, the main contribution of this work is to give a mechanism with upper bounds on the convergence rate, i.e., the relationship between the size of the market, the approximation factor, and the complexity of the class of reasonable offers.

Our contributions: We present a general framework for reducing problems of incentive-compatible mechanism design to standard algorithmic questions, for a broad class of revenue-maximizing pricing problems. To obtain our bounds we use and extend sample-complexity techniques from machine learning theory (see [3,11,30,36]) and to design our mechanisms we employ machine learning methods such as *structural risk minimization*. In general we show that an algorithm (or β-approximation) can be converted into a $(1+\epsilon)$-approximation (or $\beta(1+\epsilon)$-approximation) for the optimal mechanism design problem when the market size is at least $O(\beta\epsilon^{-2})$ times a reasonable notion of the complexity of the class of offers considered. Our formulas relating the size of the market to the approximation factor give upper bounds on the performance loss due to unknown market conditions and we view these as bounds on the *convergence rate* of our mechanism. From a learning perspective, the mechanism-design setting presents a number of technical challenges when attempting to get good bounds: in particular, the payoff function is discontinuous and asymmetric, and the payoffs for different offers are non-uniform. For example, in Section 3.3.3 we develop bounds based on a different notion of *covering number* than typically used in machine learning, in order to obtain results that are more meaningful for our setting.

We instantiate our framework for a variety of problems, some of which have been previously considered in the literature, including:

Digital Good Auction Problem: The *digital good auction problem* considers the sale of an unlimited number of units of an item to indistinguishable consumers, and has been considered by Goldberg et al. [26] and a number of subsequent papers. As argued in [26] the only reasonable offers for this setting are take-it-or-leave-it prices.

The analysis techniques developed in this paper give a *simple* proof that the random sampling auction (related to that of [26]) obtains a $(1-\epsilon)$ fraction of the optimal offer as long as the market size is at least $O(\frac{h}{\epsilon^2}\log\frac{1}{\epsilon})$ (where h is an upper bound on the valuation of any agent).

Attribute Auction Problem: The *attribute auction problem* is an abstraction of the problem using discriminatory prices based on public information (a.k.a., *attributes*) of the agents. A seller can often increase its profit by using discriminatory pricing: for example, the motion picture industry uses region encodings so that they can charge different prices for DVDs sold in different markets. Further, in many generalizations of the digital good auction problem, the agents are distinguishable via public information so the techniques exposed in the study of attribute auctions are fundamental to the study of profit maximization in general settings.

Here a reasonable class of offers to consider are mappings from the agents' attributes to take-it-or-leave-it prices. As such, we refer to these offers as *pricing functions*. For example, for one-dimensional attributes, a natural class of pricing functions might be piece-wise constant functions with k prices, as studied in [9]. In this paper we give a *general* treatment that can be applied to arbitrary classes of pricing functions. For example, if attributes are multi-dimensional, pricing functions might involve partitioning agents into markets defined by coordinate values or by some natural clustering, and then offering a constant price or a price that is some other simple function of the attributes within each market. Our bounds give a $(1+\epsilon)$-approximation when the market size is large in comparison to ϵ^{-2} scaled by a suitable notion of the complexity of the class of offers.

Combinatorial Auction Problem: We also consider the goal of profit maximization in an unlimited-supply combinatorial auction. This generalizes the digital good auction and exemplifies the problem of discriminatory pricing through the sale of multiple products. The setting here is the following. We have m different items, each in unlimited supply (like a supermarket), and bidders have valuations over *subsets* of items. Our goal is to achieve revenue nearly as large as the best revenue that uses take-it-or-leave-it prices for each item individually, i.e., the best *item-pricing*.

For arbitrary item pricings we show that our reduction has a convergence rate of $\tilde{O}\left(\frac{hm^2}{\epsilon^2}\right)$ no matter how complicated those bidders' valuations are (where the $\tilde{O}$ hides terms logarithmic in n, the number of agents; m, the number of items; and h, the highest valuation). If instead the specification of the problem constrains the item prices to be integral (e.g., in pennies) or the consumers to be *unit-demand* (desiring only one of several items) or *single-minded* (desiring only a particular bundle of items) then our bound improves to $\tilde{O}\left(\frac{hm}{\epsilon^2}\right)$. This improves on the bounds given by [21] for the unit-demand case by roughly a factor of m.

A special case of this setting is the problem of auctioning the right to traverse paths in a network. When the network is a tree and each user wants to reach the root (like drivers commuting into a city or a multicast tree in the Internet), Guruswami et al. [28] give an exact algorithm for the algorithmic problem to which our reduction applies as noted above.

Related Work: Several papers [9,10] have applied machine learning techniques to mechanism design in the context of maximizing revenue in online auctions. The online setting is more difficult than the "batch" setting we consider, but the flip-side is that as a result, that work only applies to quite simple mechanism design settings where the class $\mathcal{G}$ of allowable offers has small size and can be easily listed. Also, in a similar spirit to the goals of this paper, Awerbuch et al. [4] give reductions from online mechanism design to online optimization for a broad class of revenue maximization problems. Their work compares performance to the sum of bidders' valuations, a quite demanding measure. As a result, however, their approximation factors are necessarily logarithmic rather than $(1+\epsilon)$ as in our results.

Structure of this paper: The structure of the paper is as follows. We describe the general setting in which our results apply in Section 2 and give our generic reduction and bounds Section 3. We then apply our techniques to the digital good auction problem (Section 4), attribute auction problems (Section 5), the problem of item-pricing in combinatorial auctions (Section 6). We give our conclusions and some open research directions in Section 7.

2 Model, Notation, and Definitions

2.1 Abstract model

Our results apply to the following abstract model. We have a set of n agents $S = \{1, \ldots, n\}$. Each agent i has some private preference information v_i known only to itself (such as how much the agent is willing to pay for each of our products) and possibly also some *public* information pub_i (such as its age or location) that is known to the mechanism. A *bid* b_i is a reporting by the agent of its private information to the mechanism (which may or may not be truthful).

The basic building block of our mechanism is an *offer*. The precise notion of what an offer *is* will be defined in Section 2.2 and depends on the specific application; however, our results apply to an abstract setting where an offer g is just an incentive-compatible mechanism for a single agent that maps the agent's public information pub_i and bid b_i to a profit $g(i)$ for the mechanism. For example, in the attribute-auction problem, an offer might be a price $p(pub_i)$ that, applied to a bidder's bid b_i, either produces profit $p(pub_i)$ (and a sale) if $p(pub_i) \leq b_i$, or else produces profit 0 (and no sale) otherwise. Let $\mathcal{G}$ denote a class of offers. The specific property we assume is that if our choice of offer $g \in \mathcal{G}$ does not depend on the agent's bid b_i, then the agent will report truthfully with $b_i = v_i$.

We assume we are in an *unlimited supply* setting which in particular means that the profit from a set of bidders $S' \subseteq S$, all receiving offer g, can be written as $g(S') = \sum_{i \in S'} g(i)$.

Our approach to incentive-compatible mechanism design is via reduction to the algorithmic optimization problem. Given the *true* preferences of S and a class of offers $\mathcal{G}$, the *algorithmic optimization problem* is to find the $g \in \mathcal{G}$ with maximum profit, i.e., $\text{opt}_{\mathcal{G}} = \text{argmax}_{g \in \mathcal{G}}\, g(S)$. Let $\text{OPT}_{\mathcal{G}} = \max_{g \in \mathcal{G}} g(S)$ be this maximum profit. This computational problem is interesting in its own right; however, we will not consider it here. Instead we will assume that we are given an algorithm that either exactly solves the algorithmic optimization problem or approximates it, and our goal is to somehow use that algorithm to choose for each agent i an offer $g_i \in \mathcal{G}$ in a way that does not depend on the agent's bid b_i. Some of our techniques also make use of the existence of an algorithm that optimizes over the profit of an offer minus some penalty term that is related to the complexity of the offer, i.e., $\max_{g \in \mathcal{G}} g(S) - \mathsf{pen}_g(S)$.

One final point at this level of generality: we will assume that h is an upper bound on the value of $g(i)$ for all $i \in S$ and $g \in \mathcal{G}$; that is, no individual bidder can influence the total profit by more than h. This term will come into our general sample-complexity bounds. Auctions that make use of the technique of structural risk minimization will need to know h in advance.

2.2 Offers, Preferences, and Incentives

To describe how the framework above allows us to consider a large class of mechanism design problems, we formally discuss the details of offers, agent preferences, and the constraints imposed by incentive compatibility. To do this we develop some notation; however, the main results of the paper will be given using the general framework above.

Formally, a *market* consists of a set of n agents S and a space of possible outcomes $\mathcal{O}$. We consider *unlimited supply* allocation problems where $\mathcal{O}_i$ is set of possible outcomes (allocations) to agent i and $\mathcal{O} = \mathcal{O}_1 \times \cdots \times \mathcal{O}_n$ (i.e., all possible combinations of allocations are feasible). Except where noted, we assume there is no cost to the mechanism for producing any outcome.

As is standard in the mechanism design literature [35], an agent i's preference is fully specified by its private type, which we denote v_i. We assume *no externalities*, which means that v_i can be viewed as a preference ordering $\succeq_{v_i}$ over (outcome, payment) pairs in $\mathcal{O}_i \times \mathbb{R}$. That is, each agent cares only about what it recieves and pays, and not about what other agents get. A *bid* b_i is a reporting of one's type, i.e., it is also a preference ordering over (outcome, payment) pairs, and we say a bidder is bidding truthfully if the preference ordering under b_i matches that given by its true type v_i. Each bidder i may also have *public* information pub_i that it cannot misreport.

A deterministic mechanism is *incentive compatible* if for all agents i and all actions of the other agents, bidding truthfully is at least as good as bidding non-truthfully. If $o_i(b_i, \mathbf{b}_{-i})$ and $p_i(b_i, \mathbf{b}_{-i})$ are the outcome and payment when agent i bids b_i and the other agents bid $\mathbf{b}_{-i}$, then incentive compatibility requires for all v_i, b_i, and $\mathbf{b}_{-i}$,

$$(o_i(v_i, \mathbf{b}_{-i}), p_i(v_i, \mathbf{b}_{-i})) \succeq_{v_i} (o_i(b_i, \mathbf{b}_{-i}), p_i(b_i, \mathbf{b}_{-i})).$$

A randomized mechanism is incentive compatible if it is a randomization over deterministic incentive compatible mechanisms.

An *offer* is a mapping from a bidder's public information to a collection of (outcome, payment) pairs. We interpret making an offer to an agent as choosing the outcome and payment for them that they most prefer according to their reported type. As a result, any *fixed* offer is by definition incentive-compatible. In fact the following more general result is easy to show:

Fact 1 *A mechanism is* incentive compatible *if the choice of which offer to make to any agent does not depend on the agent's bid.*

Because all our mechanisms are incentive compatible, we will henceforth treat the profit $g(i)$ of offer g on agent i as if it were defined in terms of the *true* types v_i rather than the bids b_i.

2.3 Quasi-linear Preferences

We will apply our general framework and analysis to a number of special cases where the agents' preferences are to maximize their *quasi-linear utility*. This is the most studied case in mechanism design literature. The type, v_i, of a quasi-linear utility maximizing agent i specifies its *valuation* for each outcome. We notate valuation of agent i for outcome $o_i \in \mathcal{O}_i$ as $v_i(o_i)$. This agent's *utility* is the difference between its valuation and the price it is required to pay. I.e., for outcome o_i and payment p_i, agent i's utility is $u_i = v_i(o_i) - p_i$. An agent prefers the outcome and payment that maximizes its utility. I.e., $v_i(o_i) - p_i \geq v_i(o'_i) - p'_i$ if and only if $(o_i, p_i) \succeq_{v_i} (o'_i, p'_i)$.

For the quasi-linear case, the incentive compatibility constraints imply for all v_i, b_i, and $\mathbf{b}_{-i}$ that,

$$v_i(o_i(v_i, \mathbf{b}_{-i})) - p_i(v_i, \mathbf{b}_{-i}) \geq v_i(o_i(b_i, \mathbf{b}_{-i})) - p_i(b_i, \mathbf{b}_{-i}).$$

Notice that in the quasi-linear setting our constraint that $g(i) \leq h$ would be implied by the condition that $v_i(o_i) \leq h$ for all $o_i \in \mathcal{O}_i$.

2.4 Examples

The following examples illustrate the relationship between the outcome of the mechanism, offers, valuations, and attributes. (The first three examples are quasi-linear, the fourth is not.)

Digital Good Auction: The digital good auction models an auction of a single item in unlimited supply to indistinguishable bidders. Here the set of possible outcomes for bidder i is $\mathcal{O}_i = \{0, 1\}$ where $o_i = 1$ represents bidder i receiving a copy of the good and $o_i = 0$ otherwise. We normalize their valuation function $v_i(0) = 0$ and use a simple shorthand notation of $v_i = v_i(1)$ as the bidders privately known valuation for receiving the good. As described in the introduction, in this setting the bidders have no public information. Here, a natural class of offers, $\mathcal{G}$, is the class of all take-it-or-leave-it prices. For bidder i with valuation v_i and offer $g_p =$ "take the good for \$$p$, or leave it" the profit is

$$g_p(i) = \begin{cases} p & \text{if } p \leq v_i \\ 0 & \text{otherwise.} \end{cases}$$

We consider the digital good auction problem in detail in Section 4.

Attribute Auctions: This is the same as the digital good setting except now each bidder i is associated a public attribute, $pub_i \in \mathcal{X}$, where $\mathcal{X}$ is the *attribute space*. We view $\mathcal{X}$ as an abstract space, but one can envision it as $\mathbb{R}^d$, for example. Let $\mathcal{P}$ be a class of pricing functions from $\mathcal{X}$ to $\mathbb{R}_+$, such as all linear functions, or all functions that partition $\mathcal{X}$ into k markets in some natural way (say, based on distance to k cluster centers) and offer a different price in each. Let $\mathcal{G}$ be the class of take-it-or-leave-it offers induced by $\mathcal{P}$. That is, if $p \in \mathcal{P}$ is a pricing function, then the offer $g_p \in \mathcal{G}$ induced by p is: "for bidder i, take the good for \$$p(pub_i)$, or leave it". The profit to the mechanism from bidder i with valuation v_i and public information pub_i is

$$g_p(i) = \begin{cases} p(pub_i) & \text{if } p(pub_i) \leq v_i, \\ 0 & \text{otherwise.} \end{cases}$$

We will give analyses for several interesting classes of pricing functions in Section 5.

Combinatorial Auctions: Here we have a set J of m distinct items, each in unlimited supply. Each consumer has a private valuation $v_i(J')$ for each bundle $J' \subseteq J$ of items, which measures how much receiving bundle J' would be worth to the consumer i (again we normalize such that $v_i(\emptyset) = 0$).

For simplicity, we assume bidders are indistinguishable, i.e., there is no public information. A natural class of offers $\mathcal{G}$ (studied in [28]) is the class of functions that assign a separate price to each item, such that the price of a bundle is just the sum of the prices of the items in it (called item pricing). For price vector $\mathbf{p} = (p_1, \ldots, p_m)$ let the offer $g_{\mathbf{p}}$ = "for bundle J', pay $\sum_{j \in J'} p_j$". The profit for bidder i on offer $g_{\mathbf{p}}$ is

$$g_{\mathbf{p}}(i) = \sum \left\{ p_j \;:\; j \in \operatorname{argmax}_{J' \subset J} \left[v_i(J') - \sum_{j' \in J'} p_{j'} \right] \right\}.$$

(If the bundle J' maximizing the bidder's utility is not unique, we define the mechanism to select the utility-maximizing bundle of greatest profit.) We discuss combinatorial auctions in Section 6.

Marginal Cost Auctions with Budgets: To illustrate an interesting model with agents in a non-quasi-linear setting consider the case each bidder i's preference is given tuple (B_i, v_i) where B_i is their budget and v_i is their value-per-unit received. Possible allocations for bidder i, $\mathcal{O}_i$, are non-negative real numbers corresponding to the number of units they receive. Assuming their total payment is less than their budget, bidder i's utility is simply $v_i o_i$ minus their payment; a bidder's utility when payments exceed their budget is negative infinity.

We assume that the seller has a fixed marginal cost c for producing a unit of the good. Consider the class of offers $\mathcal{G}$ with g_p = "pay \$$p$ per unit received". A bidder i faced with offer g_p with $p < v_i$ will maximize their utility by buying enough units to exactly exhaust their budget. The payoff to the auctioneer for this bidder i is therefor B_i less c times the number of units the bidder demands. I.e.,

$$g_p(i) = \begin{cases} B_i - cB_i/p & \text{if } p \leq v_i, \\ 0 & \text{otherwise.} \end{cases}$$

This model is quite similar to one considered by Borgs et al. [12]. Though we do not explicitly analyze this setting, it is simple to apply our generic analysis to get reasonable bounds.

3 Generic Reductions

We are interested in reducing incentive-compatible mechanism design to the (non-incentive-compatible) algorithmic optimization problem. Our reductions will be based on random sampling. Let $\mathcal{A}$ be an algorithm (exact or approximate) for the algorithmic optimization problem over $\mathcal{G}$. The simplest mechanism that we consider, which we call $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ (Random Sampling Optimal offer), is the following generalization of the random sampling digital-goods auction from [26]:

(0) Bidders commit to their preferences by submitting their bids.
(1) Randomly split the bidders into two groups S_1 and S_2 by flipping a fair coin for each bidder to determine its group.
(2) Run $\mathcal{A}$ to determine the best (or approximately best) offer $g_1 \in \mathcal{G}$ over S_1, and similarly the best (or approximately best) $g_2 \in \mathcal{G}$ over S_2.
(3) Finally, apply g_1 to all bidders in S_2 and g_2 to all bidders in S_1 using their reported bids.

We will also consider various more refined versions of $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ that discretize $\mathcal{G}$ or perform some type of *structural risk minimization* (in which case we will need to assume $\mathcal{A}$ can optimize over the modifications made to $\mathcal{G}$).

Note 1: One might think that the "leave-one-out" mechanism, where the offer made to a given bidder i is the best offer for all other bidders, i.e., $\mathrm{opt}_{\mathcal{G}}(S\setminus\{i\})$, would be a better mechanism than the random sampling mechanism above. However, as pointed out in [26,25], such a mechanism (and indeed, any symmetric deterministic mechanism) has poor worst-case revenue. Furthermore, even if bidders' valuations are independently drawn from some distribution, the leave-one-out revenue can be much less stable than $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ in that it may have a non-negligable probability of achieving revenue that is far from optimal, whereas such an event is exponentially small for $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$.[3]

Note 2: The reader will notice that in converting an algorithm for finding the best offer in $\mathcal{G}$ into an incentive-compatible mechanism, we produce a mechanism whose outcome is not simply that of a single offer applied to all consumers. For example, even in the simplest case of auctioning a digital good to indistinguishable bidders, we compare our performance to the best take-it-or-leave-it price, and yet the auction itself does not in fact offer each bidder the same price (all bidders in S_1 get the same price, and all bidders in S_2 get the same price, but those two prices may be different). In fact, Goldberg and Hartline [22] show that this sort of behavior is necessary: it is not possible for an incentive-compatible auction to approximately maximize profit and offer all the bidders the same price.

[3] For example, say we are selling just one item and the distribution over valuations is 50% probability of valuation 1 and 50% probability of valuation 2. If we have n bidders, then there is a nontrivial chance (about $1/\sqrt{n}$) that there will be the exact same number of each type ($n/2$ bidders with valuation 1 and $n/2$ bidders with valuation 2), and the mechanism will make the wrong decision on everybody. The $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ mechanism on the other hand has only an exponentially small probability of doing this poorly.

The following theorem shows that the random sampling auction incurs only a small loss in performance if the profit of the optimal offer is large in comparison to the logarithm of the number of offers we are choosing from. Later sections of this paper will focus on techniques for bounding the effective size (or complexity) of $\mathcal{G}$ that can yield even stronger guarantees.

Theorem 1 *Given the offer class $\mathcal{G}$ and a β-approximation algorithm $\mathcal{A}$ for optimizing over $\mathcal{G}$, then with probability at least $1-\delta$ the profit of $RSO_{(\mathcal{G},\mathcal{A})}$ is at least $(1-\epsilon)\mathrm{OPT}_{\mathcal{G}}/\beta$ as long as*

$$\mathrm{OPT}_{\mathcal{G}} \geq \beta \tfrac{18h}{\epsilon^2} \ln\left(\tfrac{2|\mathcal{G}|}{\delta}\right).$$

Notice that this bound holds for all ϵ and δ simultaneously as these are not parameters of the mechanism. In particular, this bound and those given by the two immediate corollaries, below, show how the approximation factor improves as a function of market size.

Corollary 2 *Given the offer class $\mathcal{G}$ and a β-approximation algorithm $\mathcal{A}$ for optimizing over $\mathcal{G}$, then with probability at least $1-\delta$, the profit of $RSO_{(\mathcal{G},\mathcal{A})}$ is at least $(1-\epsilon)\mathrm{OPT}_{\mathcal{G}}/\beta$, when $\mathrm{OPT}_{\mathcal{G}} \geq n$ and the number of bidders n satisfies*

$$n \geq \tfrac{18h\beta}{\epsilon^2} \ln\left(\tfrac{2|\mathcal{G}|}{\delta}\right).$$

Corollary 3 *Given the offer class $\mathcal{G}$ and a β-approximation algorithm $\mathcal{A}$ for optimizing over $\mathcal{G}$ then with probability at least $1-\delta$, the profit of $RSO_{(\mathcal{G},\mathcal{A})}$ is at least*

$$(1-\epsilon)\mathrm{OPT}_{\mathcal{G}}/\beta - \tfrac{18h\beta}{\epsilon^2} \ln\left(\tfrac{2|\mathcal{G}|}{\delta}\right).$$

If bidders' valuations are in the interval $[1,h]$ and the take-it-or-leave-it offer of \$1 is in $\mathcal{G}$, then the condition $\mathrm{OPT}_{\mathcal{G}} \geq n$ is trivially satisfied and Corollary 2 can be interpreted as giving a bound on the *convergence rate* of the random sampling auction. Corollary 3 is a useful form of our bound when considering structural risk minimization and it also matches the form of bounds given in prior work (e.g., [9]).

For example, in the digital good auction with the class of offers $\mathcal{G}_\epsilon$ consisting of all take-it-or-leave-it offers in the interval $[1,h]$ discretized to powers of $1+\epsilon$, we have $\mathrm{OPT}_{\mathcal{G}_\epsilon} \geq n$ (since each bidder's valuation is at least 1), $\beta = 1$ (since the algorithmic problem is easy), and $|\mathcal{G}_\epsilon| = \lceil \log_{1+\epsilon} h \rceil$. So, Corollary 2 states that $O(\frac{h}{\epsilon^2} \log \log_{1+\epsilon} h)$ bidders are sufficient to perform nearly as well as optimal (we derive better bounds for this problem in Section 4).

In general we will give our bounds in a similar form as Theorem 1, knowing that bounds of the form of Corollary 2 and 3 can be easily derived. The only exceptions are the structural risk minimization results which we give in the same form as Corollary 3.

In the remainder of this section we prove Theorem 1. We start with a lemma that is key to our analysis.

Lemma 4 *Given S, an offer g satisfying $0 \leq g(i) \leq h$ for all $i \in S$, and a profit level p, if we randomly partition S into S_1 and S_2, then the probability that $|g(S_1) - g(S_2)| \geq \epsilon \max[g(S), p]$ is at most $2e^{\left[-\frac{\epsilon^2 p}{2h}\right]}$.*

Proof: Let $Y_1, \ldots, Y_n$ be i.i.d. random variables that define the partition of S into S_1 and S_2: that is, Y_i is 1 with probability $\frac{1}{2}$ and Y_i is 2 with probability $\frac{1}{2}$. Let $t(Y_1, ..., Y_n) = \sum_{i:Y_i=1} g(i)$. So, as a random variable, $g(S_1) = t(Y_1, ..., Y_n)$ and clearly $\mathbf{E}[t(Y_1, ..., Y_n)] = \frac{g(S)}{2}$. Assume first that $g(S) \geq p$. From the McDiarmid concentration inequality (see Theorem 26 in Appendix A), by plugging in $c_i = g(i)$, we get:

$$\Pr\left\{\left|g(S_1) - \frac{g(S)}{2}\right| \geq \frac{\epsilon}{2} g(S)\right\} \leq 2e^{-\frac{1}{2}\epsilon^2 g(S)^2 / \sum_{i=1}^{n} g(i)^2}.$$

Since

$$\sum_{i=1}^{n} g(i)^2 \leq \max_i \{g(i)\} \sum_{i=1}^{n} g(i) \leq h g(S),$$

we obtain:

$$\Pr\left\{\left|g(S_1) - \frac{g(S)}{2}\right| \geq \frac{\epsilon}{2} g(S)\right\} \leq 2e^{-\left[\frac{\epsilon^2 g(S)}{2h}\right]}.$$

Moreover, since $g(S_1) + g(S_2) = g(S)$ and $g(S) \geq p$, we obtain $\Pr\{|g(S_1) - g(S_2)| \geq \epsilon g(S)\} \leq 2e^{-\epsilon^2 p/(2h)}$, as desired. Consider now the case that $g(S) < p$. Again, using the McDiarmid inequality we have

$$\Pr\{|g(S_1) - g(S_2)| \geq \epsilon p\} \leq 2e^{-\frac{1}{2}\epsilon^2 p^2 / \sum_{i=1}^{n} g(i)^2}.$$

Since $\sum_{i=1}^{n} g(i)^2 \leq h g(S) \leq ph$ we obtain again that

$$\Pr\{|g(S_1) - g(S_2)| \geq \epsilon p\} \leq 2e^{\left[-\frac{\epsilon^2 p}{2h}\right]},$$

which gives us the desired bound. □

It is worth noting that using tail inequalities that depend on the maximum range of the random variables rather than the sum of their squares in the proof of Lemma 4 would increase the h to an h^2 in the exponent. Note also that if $g(i) = g'(i)$ for all $i \in S$ then they are equivalent from the point of view of the auction; we will use $|\mathcal{G}|$ to denote the number of *different* such offers in $\mathcal{G}$.[4] Lemma 4 implies that:

Corollary 5 *For a random partition of S into S_1 and S_2, with probability at least $1-\delta$, all offers g in $\mathcal{G}$ such that $g(S) \geq \frac{2h}{\epsilon^2}\ln\left(\frac{2|\mathcal{G}|}{\delta}\right)$ satisfy $|g(S_1) - g(S_2)| \leq \epsilon g(S)$.*

Proof: Follows from Lemma 4 by plugging in $p = \frac{2h}{\epsilon^2}\ln\left(\frac{2|\mathcal{G}|}{\delta}\right)$ and then using the union bound over all $g \in \mathcal{G}$. □

We complete this section with the proof of the main theorem.

Proof of of Theorem 1: Let g_1 be the offer in $\mathcal{G}$ produced by $\mathcal{A}$ over S_1 and g_2 be the offer in $\mathcal{G}$ produced by $\mathcal{A}$ over S_2. Let g_{OPT} be the optimal offer in $\mathcal{G}$ over S; so $g_{\mathrm{OPT}}(S) = \mathrm{OPT}_{\mathcal{G}}$. Since the optimal offer over S_1 is at least as good as g_{OPT} on S_1 (and likewise for S_2), the fact that $\mathcal{A}$ is a β-approximation implies that $g_1(S_1) \geq \frac{g_{\mathrm{OPT}}(S_1)}{\beta}$ and $g_2(S_2) \geq \frac{g_{\mathrm{OPT}}(S_2)}{\beta}$.

Let $p = \frac{18h}{\epsilon^2}\ln\left(\frac{2|\mathcal{G}|}{\delta}\right)$. Using Lemma 4 (applying the union bound over all $g \in \mathcal{G}$), we have that with probability $1-\delta$, every $g \in \mathcal{G}$ satisfies $|g(S_1) - g(S_2)| \leq \frac{\epsilon}{3}\max[g(S), p]$. In particular, $g_1(S_2) \geq g_1(S_1) - \frac{\epsilon}{3}\max[g_1(S), p]$, and $g_2(S_1) \geq g_2(S_2) - \frac{\epsilon}{3}\max[g_2(S), p]$.

Since the theorem assumes that $\mathrm{OPT}_{\mathcal{G}} \geq \beta p$, summing the above two inequalities and performing a case analysis[5] we get that the profit of $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$, namely the sum $g_1(S_2) + g_2(S_1)$, is at least $(1-\epsilon)\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta}$. More specifically, assume first that $g_1(S) \geq p$ and $g_2(S) \geq p$. This implies that $g_1(S_2) \geq g_1(S_1) - \frac{\epsilon}{3}g_1(S)$ and $g_2(S_1) \geq g_2(S_2) - \frac{\epsilon}{3}g_2(S)$, and therefore $(1+\frac{\epsilon}{3})g_1(S_2) \geq (1-\frac{\epsilon}{3})g_1(S_1)$ and $(1+\frac{\epsilon}{3})g_2(S_1) \geq (1-\frac{\epsilon}{3})g_2(S_2)$. So, the profit of $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ in this case is at least

$$\frac{1-\frac{\epsilon}{3}}{1+\frac{\epsilon}{3}}(g_1(S_1) + g_2(S_2)) \geq \frac{1-\frac{\epsilon}{3}}{1+\frac{\epsilon}{3}}\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta} \geq (1-\epsilon)\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta}.$$

If both $g_1(S) < p$ and $g_2(S) < p$, then $g_1(S_2) \geq g_1(S_1) - \frac{\epsilon}{3}p$ and $g_2(S_1) \geq g_2(S_2) - \frac{\epsilon}{3}p$, and so the profit of $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ in this case is at least $\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta} - \frac{2\epsilon}{3}p$ which is at least $(1-\epsilon)\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta}$ by our assumption that $\mathrm{OPT}_{\mathcal{G}} \geq \beta p$. Finally, assume without loss of generality that $g_1(S) \geq p$ and $g_2(S) < p$. This implies that $g_1(S_2) \geq g_1(S_1) - \frac{\epsilon}{3}g_1(S)$ and $g_2(S_1) \geq g_2(S_2) - \frac{\epsilon}{3}p$. The former inequality implies that $(1+\frac{\epsilon}{3})g_1(S_2) \geq (1-\frac{\epsilon}{3})g_1(S_1)$, and so $g_1(S_2) \geq \left(1 - \frac{2\epsilon}{3}\right)g_1(S_1)$,

[4] Notice that in our generic reduction, $|\mathcal{G}|$ only appears in the analysis and we do not actually have to know whether two offers are equivalent with respect to S when running the auction.

[5] Note that if $\beta = 1$, then the conclusion follows easily. The case analysis is only need to deal with the case $\beta > 1$.

and the latter inequality implies that $g_2(S_1) \geq g_2(S_2) - \frac{\epsilon}{3}\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta}$. Together we have that

$$g_1(S_2) + g_2(S_1) \geq \left(1 - \frac{2\epsilon}{3}\right)\frac{g_{\mathrm{OPT}}(S_1)}{\beta} + \frac{g_{\mathrm{OPT}}(S_2)}{\beta} - \frac{\epsilon}{3}\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta} \geq (1-\epsilon)\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta},$$

as desired. □

3.2 Structural Risk Minimization

In many natural cases, $\mathcal{G}$ consists of offers at different "levels of complexity" k. In the case of attribute auctions, for instance, $\mathcal{G}$ could be an offer class induced by pricing functions that partition bidders into k markets and offer a constant price in each market, for different values of k. The larger k is the more complex the offer is. One natural approach to such a setting is to perform *structural risk minimization* (SRM): that is, to assign a penalty term to offers based on their complexity and then to run a version of $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ in which $\mathcal{A}$ optimizes profit minus penalty. Specifically, let $\bar{\mathcal{G}}$ be a series of offers classes $\mathcal{G}_1, \mathcal{G}_2, \ldots$, and let pen be a penalty function defined over these classes. We then define the procedure $\mathrm{RSO\text{-}SRM}_{(\bar{\mathcal{G}},\mathsf{pen})}$ as follows:

1. Randomly partition the bidders into two sets, S_1 and S_2, by flipping fair coin for each bidder.
2. Compute g_1 to maximize $\max_k \max_{g\in\mathcal{G}_k}[g(S_1) - \mathsf{pen}(\mathcal{G}_k)]$ and similarly compute g_2 from S_2.
3. Use the offer g_1 for bidders in S_2 and the offer g_2 for bidders in S_1.

We can now derive a guarantee for the $\mathrm{RSO\text{-}SRM}_{(\bar{\mathcal{G}},\mathsf{pen})}$ mechanism as follows:

Theorem 6 *Assuming that we have an algorithm for solving the optimization problem required by RSO-SRM*$_{(\bar{\mathcal{G}},\mathsf{pen})}$*, then for any given value of* n, ϵ*, and* δ*, with probability at least* $1-\delta$*, the revenue of RSO-SRM*$_{(\bar{\mathcal{G}},\mathsf{pen})}$ *for* $\mathsf{pen}(\mathcal{G}_k) = \frac{8h_k}{\epsilon^2}\ln\left(\frac{8k^2|\mathcal{G}_k|}{\delta}\right)$ *is at least*

$$\max_k\left([(1-\epsilon)\,\mathrm{OPT}_k - 2\mathsf{pen}(\mathcal{G}_k)]\right),$$

where h_k *is the maximum payoff from* $\mathcal{G}_k$ *and* $\mathrm{OPT}_k = \mathrm{OPT}_{\mathcal{G}_k}$.

Proof: Using Corollary 5 and a union bound over the values $\delta_k = \delta/(4k^2)$, we obtain that with probability at least $1-\delta$, simultaneously for all k and for all offers g in $\mathcal{G}_k$ such that $g(S) \geq \frac{8h_k}{\epsilon^2}\ln(8k^2|\mathcal{G}_k|/\delta) = \mathsf{pen}(\mathcal{G}_k)$, we have $|g(S_1) - g(S_2)| \leq \frac{\epsilon}{2}g(S)$. Let k^* be the optimal index, namely let k^* be the index such that $(1-\epsilon)\,\mathrm{OPT}_{k^*} - 2\mathsf{pen}(\mathcal{G}_{k^*}) = \max_k((1-\epsilon)\,\mathrm{OPT}_k - 2\mathsf{pen}(\mathcal{G}_k))$, and let k_i be the index of the best offer (according to our criterion) over S_i, for $i = 1, 2$. By our assumption that g_1 and g_2 were chosen by an optimal algorithm, we have $g_i(S_i) - \mathsf{pen}(\mathcal{G}_{k_i}) \geq g_{\mathrm{OPT}_{k^*}}(S_i) - \mathsf{pen}(\mathcal{G}_{k^*})$, for $i = 1, 2$.

We will argue next that $g_1(S_2) \geq \frac{1-\frac{\epsilon}{2}}{1+\frac{\epsilon}{2}} (g_{\mathrm{OPT}_{k^*}}(S_1) - \mathsf{pen}(\mathcal{G}_{k^*}))$. First, if $g_1(S_1) < \mathsf{pen}(\mathcal{G}_{k_1})$, then the conclusion is clear since we have $0 > g_1(S_1) - \mathsf{pen}(\mathcal{G}_{k_1}) \geq g_{\mathrm{OPT}_{k^*}}(S_1) - \mathsf{pen}(\mathcal{G}_{k^*})$. If $g_1(S_1) \geq \mathsf{pen}(\mathcal{G}_{k_1})$, then as argued above we have $|g_1(S_1) - g_1(S_2)| \leq \frac{\epsilon}{2} g_1(S)$ and so

$$g_1(S_2) \geq \frac{1-\frac{\epsilon}{2}}{1+\frac{\epsilon}{2}} g_1(S_1) \geq \frac{1-\frac{\epsilon}{2}}{1+\frac{\epsilon}{2}} \left(g_{\mathrm{OPT}_{k^*}}(S_1) - \mathsf{pen}(\mathcal{G}_{k^*})\right).$$

Similarly, we can prove that we have $g_2(S_1) \geq \frac{1-\frac{\epsilon}{2}}{1+\frac{\epsilon}{2}} (g_{\mathrm{OPT}_{k^*}}(S_2) - \mathsf{pen}(\mathcal{G}_{k^*}))$. All these together imply that the profit of the mechanism $\mathrm{RSO\text{-}SRM}_{(\bar{\mathcal{G}},\mathsf{pen})}$, namely $g_1(S_2) + g_2(S_1)$, is at least

$$\frac{1-\frac{\epsilon}{2}}{1+\frac{\epsilon}{2}} \left(g_{\mathrm{OPT}_{k^*}}(S) - 2\mathsf{pen}(\mathcal{G}_{k^*})\right) \geq \left((1-\epsilon)\,\mathrm{OPT}_{k^*} - 2\mathsf{pen}(\mathcal{G}_{k^*})\right),$$

as desired. □

3.3 Improving the Bounds

The results above say, in essence, that if we have enough bidders so that the optimal profit is large compared to $\frac{h}{\epsilon^2}\log(|\mathcal{G}|)$, then our mechanism will perform nearly as well as the best offer in $\mathcal{G}$. In these bounds, one should think of $\log(|\mathcal{G}|)$ as a measure of the complexity of the offer class $\mathcal{G}$; for instance, it can be thought of as the number of bits needed to describe a typical offer in that class. However, in many cases one can achieve a better bound by adapting techniques developed for analyzing generalization performance in machine learning theory. In this section, we discuss a number of such methods that can produce better bounds. These include both *analysis* techniques (such as using appropriate forms of *covering numbers*), where we do not change the mechanism but instead provide a stronger guarantee, and *design* techniques (like *discretizing*), where we modify the mechanism to produce a better bound.

3.3.1 Discretizing

Notation: Given a class of offers $\mathcal{G}$, define $\mathcal{G}_\alpha$ to be the set of offers induced by rounding all prices down to the nearest power of $(1+\alpha)$.

In many cases, we can greatly reduce $|\mathcal{G}|$ without much affecting $\mathrm{OPT}_{\mathcal{G}}$ by performing some type of discretization. For instance, for auctioning a digital good, there are infinitely many offers induced by all take-it-or-leave-it prices but only $\log_{1+\alpha} h \approx \frac{1}{\alpha} \ln h$ offers induced by the discretized prices at powers of $1+\alpha$. Also, since rounding down the optimal price to the nearest power of $1+\alpha$ can reduce revenue for this auction by at most a factor of $1+\alpha$, the optimal offer in the discretized class must be close, in terms of total profit, to the optimal offer in the original class. More generally, if we can find a smaller offer class $\mathcal{G}'$ such that $\mathrm{OPT}_{\mathcal{G}'}$ is guaranteed to be close to $\mathrm{OPT}_{\mathcal{G}}$, then we can instruct our algorithm $\mathcal{A}$ to optimize over $\mathcal{G}'$ instead of $\mathcal{G}$ to get better bounds. We consider the discretization $\mathcal{G}_\alpha$ in our refined analysis of the digital good auction problem (Section 4) and in our consideration of attribute auctions (Section 5). Further, in Section 6 we discuss an interesting alternative discretization for item-pricing in combinatorial auctions.

3.3.2 Counting Possible Outputs

Suppose we can argue that our algorithm $\mathcal{A}$, run on a subset of S, will only ever output offers from a restricted set $\mathcal{G}_\mathcal{A} \subseteq \mathcal{G}$. For example, for the problem of auctioning a digital good, if $\mathcal{A}$ picks the offer based on the optimal take-it-or-leave-it price over its input then this price must be one of the bids, so $|\mathcal{G}_\mathcal{A}| \leq n$. Then, we can simply replace $|\mathcal{G}|$ with $|\mathcal{G}_\mathcal{A}|$ (or $|\mathcal{G}_\mathcal{A}|+1$ if the optimal offer is not in $\mathcal{G}_\mathcal{A}$) in all the above arguments. Formally we can say that:

Observation 7 *If algorithm $\mathcal{A}$, run on any subset of S, only output offers from a restricted set $\mathcal{G}_\mathcal{A} \subseteq \mathcal{G}$, then all the bounds in Sections 3.1 and 3.2 hold with $|\mathcal{G}|$ replaced by $|\mathcal{G}_\mathcal{A}|+1$.*

3.3.3 Using Covering Numbers

The main idea of these arguments is the following. Suppose $\mathcal{G}$ has the property that there exists a much smaller class $\mathcal{G}'$ such that every $g \in \mathcal{G}$ is "close" to some $g' \in \mathcal{G}'$, with respect to the given set of bidders S. Then one can show that if all offers in $\mathcal{G}'$ perform similarly on S_1 as they do on S_2, then this will be true for all offers in $\mathcal{G}$ as well. These kind of arguments are quite often used in machine learning (see for instance [3,13,16,36]), but the main challenge is to define the right notion of "close" for our mechanism design setting to get good and meaningful bounds. Specifically, we will consider L_1 multiplicative γ-covers which we define as follows:

Definition 1 *$\mathcal{G}'$ is an L_1 multiplicative γ-cover of $\mathcal{G}$ with respect to S if for every $g \in \mathcal{G}$ there exists $g' \in \mathcal{G}'$ such that*

$$\sum_{i \in S} |g(i) - g'(i)| \;\; \leq \;\; \gamma g(S).$$

In the following we present bounds based on L_1 multiplicative γ-covers. We start by proving the following structural lemma characterizing these L_1 covers.

Lemma 8 *If* $\sum_{i \in S} |g(i) - g'(i)| \leq \gamma g(S)$ *and* $|g'(S_1) - g'(S_2)| \leq \epsilon' \max[g'(S), p]$ *then we have* $|g(S_1) - g(S_2)| \leq \epsilon' \max[g'(S), p] + \gamma g(S)$*. This further implies that* $|g(S_1) - g(S_2)| \leq (\gamma + \epsilon'(1+\gamma)) \max[g(S), p]$*.*

Proof: We will first prove that $g(S_1) \geq g(S_2) - \epsilon' \max[g'(S), p] - \gamma g(S)$. Note that this clearly implies $g(S_1) \geq g(S_2) - (\gamma + \epsilon'(1+\gamma)) \max[g(S), p]$, since the first assumption in the lemma implies that $|g(S) - g'(S)| \leq \gamma g(S)$. Let us define $\vec{\Delta}_{g_1 g_2}(S) = \sum_{i \in S} \max(g_1(i) - g_2(i), 0)$ and consider $\Delta_{gg'}(S) = \vec{\Delta}_{gg'}(S) + \vec{\Delta}_{g'g}(S) = \sum_{i \in S} |g(i) - g'(i)|$. Clearly, for any $S' \subseteq S$ we have $\vec{\Delta}_{gg'}(S) \geq \vec{\Delta}_{gg'}(S')$ and likewise $\Delta_{gg'}(S) \geq \Delta_{gg'}(S')$. Also, for any subset $S' \subseteq S$ we have $g(S') - g'(S') \leq \vec{\Delta}_{gg'}(S)$ and $g'(S') - g(S') \leq \vec{\Delta}_{g'g}(S)$. Now, from $g'(S_1) \geq g'(S_2) - \epsilon' \max[g'(S), p]$ we obtain that $g(S_1) + \vec{\Delta}_{g'g}(S) \geq g'(S_2) - \epsilon' \max[g'(S), p] \geq g(S_2) - \vec{\Delta}_{gg'}(S) - \epsilon' \max[g'(S), p]$. Therefore we have $g(S_1) \geq g(S_2) - \Delta_{gg'}(S) - \epsilon' \max[g'(S), p]$, which implies $g(S_1) \geq g(S_2) - \epsilon' \max[g'(S), p] - \gamma g(S)$, as desired. Using the same argument with S_1 replaced by S_2 yields the theorem. $\square$

Using Lemma 8, we can now get the following bound:

Theorem 9 *Given the offer class* $\mathcal{G}$ *and a* β*-approximation algorithm* $\mathcal{A}$ *for optimizing over* $\mathcal{G}$*, then with probability at least* $1 - \delta$*, the profit of* $RSO_{(\mathcal{G},\mathcal{A})}$ *is at least* $(1-\epsilon)\mathrm{OPT}_{\mathcal{G}}/\beta$ *so long as*

$$\mathrm{OPT}_{\mathcal{G}} \geq \beta \tfrac{72h}{\epsilon^2} \ln\left(\tfrac{2|\mathcal{G}'|}{\delta}\right),$$

for some L_1 *multiplicative* $\frac{\epsilon}{12}$*-cover* $\mathcal{G}'$ *of* $\mathcal{G}$ *with respect to* S*.*

Proof: Let $p = \frac{72h}{\epsilon^2} \ln\left(\frac{2|\mathcal{G}'|}{\delta}\right)$. By Lemma 4, applying the union bound, we have that with probability $1 - \delta$, every $g' \in \mathcal{G}'$ satisfies $|g'(S_1) - g'(S_2)| \leq \frac{\epsilon}{6} \max[g'(S), p]$. Using Lemma 8, with ϵ' set to $\frac{\epsilon}{6}$ and γ set to $\frac{\epsilon}{12}$, we obtain that with probability $1 - \delta$, every $g \in \mathcal{G}$ satisfies $|g(S_1) - g(S_2)| \leq \frac{\epsilon}{3} \max[g(S), p]$. Finally, proceeding as in the proof of Theorem 1 we obtain the desired result. $\square$

Notice that Theorem 9 implies that:

Corollary 10 *Given the offer class* $\mathcal{G}$ *and a* β*-approximation algorithm* $\mathcal{A}$ *for optimizing over* $\mathcal{G}$*, then with probability at least* $1 - \delta$*, the profit of* $RSO_{(\mathcal{G},\mathcal{A})}$ *is at least* $(1-\epsilon)\mathrm{OPT}_{\mathcal{G}}/\beta$*, so long as* $\mathrm{OPT}_{\mathcal{G}} \geq n$ *and the number of bidders satisfies*

$$n \geq \tfrac{72h\beta}{\epsilon^2} \ln\left(\tfrac{2|\mathcal{G}'|}{\delta}\right)$$

for some L_1 *multiplicative* $\frac{\epsilon}{12}$*-cover* $\mathcal{G}'$ *of* $\mathcal{G}$ *with respect to* S*.*

We will demonstrate the utility of L_1 multiplicative covers in Section 4 by showing the existence of L_1 covers of size $o(n)$ for the digital good auction. It is worth noting that a straightforward application of analogous ϵ-cover results in learning theory [3] (which would require an additive, rather than multiplicative gap of ϵ for every bidder) would add an extra factor of h into our sample-size bounds.

4 The Digital Good Auction

We now consider applying the results in Section 3 to the problem of auctioning a digital good to indistinguishable bidders. In this section we define $\mathcal{G}$ to be the natural class of offers induced by the set of all take-it-or-leave-it prices (see for instance [25]). Clearly in this case, it is trivial to solve the underlying optimization problem optimally: given a set of bidders, just output the offer induced by the constant price that maximizes the price times the number of bidders with bids at least as high as the price. Also, it is easy to see that this price will be one of the bid values. Thus, applying Theorem 7 with the bound on $|\mathcal{G}_\mathcal{A}| = n$, we get an approximately optimal auction with convergence rate $O(h \log n)$.

We can obtain better results using L_1 multiplicative-cover arguments and Theorem 9 as follows. Let $b_1, \ldots, b_n$ be the bids of the n bidders sorted from highest to lowest. Define $\mathcal{G}'$ as the offer class induced by $\{b_i \ : \ i = \lfloor (1+\gamma)^j \rfloor \text{ for some } j \in \mathbb{Z}\} \cup \{(1+\gamma)^i \ : \ i \in \{1, \ldots, \log_{1+\gamma} h\}\}$. Consider $g \in \mathcal{G}$ and find the $g' \in \mathcal{G}'$ that offers the largest price less than the offer price of g. Notice first that all the winners in S on g also win in g'. Second, the offer price of g' is within a factor of $1+\gamma$ of the offer price of g. Third, g' has at most a factor of $1+\gamma$ more winners than g. The first two facts above imply that $\vec{\Delta}_{gg'}(S) \leq \gamma g(S)$. The third fact implies that $\vec{\Delta}_{g'g}(S) \leq \gamma g(S)$. Thus, $\Delta_{gg'} \leq 2\gamma g(S)$ and therefore, $\mathcal{G}'$ is a 2γ-cover of $\mathcal{G}$ (see the proof of Lemma 8 for definitions of $\Delta_{gg'}$ and $\vec{\Delta}_{gg'}$). Since $|\mathcal{G}'|$ is $O(\log hn)$, the additive loss of $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ is $O(h \log \log nh)$. [6]

We can also apply the discretization technique by defining $\mathcal{G}_\alpha$ to be the set of offers induced by the set of all constant-price functions whose price $v \in [1, h]$ is a power of $(1+\alpha)$ and $\alpha = \frac{\epsilon}{2}$. Clearly, if we can get revenue at least $(1 - \frac{\epsilon}{2})$ times the optimal in this class, we will be within $(1-\epsilon)$ of the optimal fixed price overall. For example, Corollary 2 ($\mathcal{A}$ can trivially find the best offer in

[6] It is interesting to contrast these results with that of [26] which showed that RSO over the set of constant-price functions is near 6-competitive with the promise that $n \gg h$. A much more complicated analysis of RSO in a slightly different competitive framework is given in [18].

$\mathcal{G}'$ by simply trying all of them) shows that with probability $1-\delta$ we get at least $1-\epsilon$ times the revenue of the optimal take-it-or-leave-it offer so long as the number of bidders n is at least $\frac{72h}{\epsilon^2}\ln(\frac{4\ln h}{\epsilon\delta}) = O(h\log\log h)$.

4.1 Data Dependent Bounds

We can use the high level idea of our structural risk minimization reduction in order to get a better *data dependent* bound for the digital good auction. In particular, we can replace the "h" term in the additive loss with the actual sale price used by the optimal take-it-or-leave-it offer (in fact, even better, the lowest sales price needed to generate near-optimal revenue), yielding a much better bound when most of the profit to be made is from the low bids. The idea is that rather than penalizing the "complexity" of the offer in the usual sense, we instead penalize the use of higher prices.

Let $q_i = (1+\alpha)^i$ and offer g_i be the take-it-or-leave-it price of q_i. Define $\bar{\mathcal{G}} = \{g_1\}, \{g_2\}, \ldots$ and consider the auction RSO-SRM$_{\bar{\mathcal{G}},\mathsf{pen}}$ with $\mathsf{pen}(\{g_i\})$ specified from Section 3.2 to be $\frac{8q_i}{\epsilon^2}\ln\left(\frac{8i^2}{\delta}\right)$. The following is an a corollary of of Theorem 6.

Corollary 11 *For any given value of n, ϵ, and δ, with probability $1-\delta$, the revenue of RSO-SRM$_{(\bar{\mathcal{G}},\mathsf{pen})}$ is at least $\max_i\left[(1-\epsilon)g_i(S) - 2\mathsf{pen}(\{g_i\})\right]$, where $\mathsf{pen}(\{g_i\}) = \frac{8q_i}{\epsilon^2}\ln\left(\frac{8i^2}{\delta}\right)$.*

In other words, if the optimal take-it-or-leave-it offer has a sale price of p, then RSO-SRM$_{(\bar{\mathcal{G}},\mathsf{pen})}$ has convergence rate bounded by $O(p\log\log h)$ instead of $O(h\log\log h)$ as provided by our generic analysis of RSO$_{(\mathcal{G},\mathcal{A})}$.

4.2 A Special Purpose Analysis for the Digital Good Auction

In this section we present a refined data independent analysis for the digital good auction. Specifically, we can show for an optimal algorithm $\mathcal{A}$, that:

Theorem 12 *For $\delta < \frac{1}{2}$, with probability $1-\delta$, RSO$_{(\mathcal{G}_\alpha,\mathcal{A})}$ obtains profit at least*

$$\mathrm{OPT}_{\mathcal{G}_\alpha} - 8\sqrt{h\,\mathrm{OPT}_{\mathcal{G}_\alpha}\log\left(\frac{1}{\alpha\delta}\right)}.$$

Corollary 13 *For $\delta < \frac{1}{2}$ and $\alpha = \frac{\epsilon}{2}$, so long as $\mathrm{OPT}_{\mathcal{G}_\alpha} \geq (\frac{16}{\epsilon})^2 h\log\left(\frac{2}{\epsilon\delta}\right)$, then with probability at least $1-\delta$, the profit of RSO$_{(\mathcal{G}_\alpha,\mathcal{A})}$ is at least $(1-\epsilon)\,\mathrm{OPT}_{\mathcal{G}}$.*

The above corollary improves over our basic discretization results using Theorem 1 by an $O(\log\log h)$ factor in the convergence rate.

To prove Theorem 12, let us introduce some notation. For the offer g_v induced by the take-it-or-leave-it offer of price v, let n_v denote the number of winners (bidders whose value is at least v), and let $r_v = v \cdot n_v$ denote the profit of g_v on S. Denote by $\hat{r}_v$ the observed profit of g_v on S_1 (and so $\hat{r}_v = v \cdot \hat{n}_v$, where $\hat{n}_v$ is the number of winners in S_1 for g_v). So, we have $\mathbf{E}[\hat{r}_v] = \frac{r_v}{2}$. We now begin with the following lemma.

Lemma 14 *Let $\epsilon < 1$ and $\delta < \frac{1}{2}$. With probability at least $1-\delta$ we have that, for every $g_v \in \mathcal{G}_\alpha$ the observed profit on S_1 satisfies:*

$$\left|\hat{r}_v - \frac{r_v}{2}\right| \leq \max\left(\frac{h\log\left(\frac{1}{\alpha\delta}\right)}{\epsilon}, \epsilon r_v\right).$$

Proof: First for a given price v let $a_{n,v}$ be $|\hat{n}_v - \frac{n_v}{2}|$. To prove our lemma we will use the consequence of Chernoff bound we present in Appendix A, Theorem 27. For any v and $j \geq 1$ we consider $n' = \frac{(1+\alpha)^j \log\left(\frac{1}{\alpha\delta}\right)}{\epsilon^2}$, and so we get

$$\Pr\left\{a_{n,v} \geq \epsilon \max\left(n_v, \frac{(1+\alpha)^j \log\left(\frac{1}{\alpha\delta}\right)}{\epsilon^2}\right)\right\} \leq 2e^{-2(1+\alpha)^j \log\left(\frac{1}{\alpha\delta}\right)}.$$

This further implies that we have $a_{n,v} \geq \epsilon \max\left(n_v, \frac{(1+\alpha)^j \log\left(\frac{1}{\alpha\delta}\right)}{\epsilon^2}\right)$ with probability at most $2(\alpha\delta)^{2(1+\alpha)^j}$. Therefore for $v = \frac{h}{(1+\alpha)^j}$ we have

$$\Pr\left\{\left|\hat{r}_v - \frac{r_v}{2}\right| \geq \max\left(\frac{h\log\left(\frac{1}{\alpha\delta}\right)}{\epsilon}, \epsilon r_v\right)\right\} \leq 2(\alpha\delta)^{2(1+\alpha)^j},$$

and so the probability that there exists a $g_v \in \mathcal{G}_\alpha$ such that $\left|\hat{r}_v - \frac{r_v}{2}\right| \geq \max\left(\frac{h}{\epsilon}, \epsilon r_v\right)$ is at most $2\sum_j (\alpha\delta)^{2(1+\alpha)^j} \leq 2\sum_{j'} \frac{1}{\alpha}(\alpha\delta)^{2\cdot 2^{j'}} \leq \delta$. This implies that with high probability, at least $1-\delta$, we have that simultaneously, for every $g_v \in \mathcal{G}_\alpha$ the observed revenue on S_1 satisfies:

$$\left|\hat{r}_v - \frac{r_v}{2}\right| \leq \max\left(\frac{h\log\left(\frac{1}{\alpha\delta}\right)}{\epsilon}, \epsilon r_v\right),$$

as desired. □

Proof of Theorem 12: Assume now that it is the case that for every $g_v \in \mathcal{G}_\alpha$ we have $\left|\hat{r}_v - \frac{r_v}{2}\right| \leq \max\left(\frac{H}{\epsilon}, \epsilon r_v\right)$, where $H = h\log\left(\frac{1}{\alpha\delta}\right)$. Let v^* be the optimal price level among prices in $\mathcal{G}_\alpha$, and let $\tilde{v}^*$ be the price that looks best on S_1. Obviously, our gain on S_2 is $r_{\tilde{v}^*} - \hat{r}_{\tilde{v}^*}$. We have $\hat{r}_{v^*} \geq \frac{r_v^*}{2} - \frac{H}{\epsilon} - \epsilon r_{v^*} r_{v^*} \frac{1-2\epsilon}{2} - \frac{H}{\epsilon}$, $\hat{r}_{\tilde{v}^*} \geq \hat{r}_{v^*}$ and $\hat{r}_{\tilde{v}^*} \leq \frac{r_{\tilde{v}^*}}{2} + \frac{H}{\epsilon} + \epsilon r_{\tilde{v}^*} \leq \frac{r_{\tilde{v}^*}}{2} + \frac{H}{\epsilon} + \epsilon r_{v^*}$, and therefore $r_{\tilde{v}^*} - \hat{r}_{\tilde{v}^*} \geq \hat{r}_{\tilde{v}^*} - \frac{H}{\epsilon} - \epsilon r_{v^*}$, which finally implies that $r_{\tilde{v}^*} - \hat{r}_{\tilde{v}^*} \geq r_{v^*}\left(\frac{1}{2} - 2\epsilon\right) - 2\frac{H}{\epsilon}$. This implies that with probability at least $1 - \frac{\delta}{2}$ our gain on S_2 is at least $r_{v^*}\left(\frac{1}{2} - 2\epsilon\right) - 2\frac{H}{\epsilon}$, and similarly our gain on S_1 is at least $r_{v^*}\left(\frac{1}{2} - 2\epsilon\right) - 2\frac{H}{\epsilon}$. Therefore, with probability $1-\delta$, our revenue is $\mathrm{OPT}_{\mathcal{G}_\alpha}(1-4\epsilon) - 4\frac{h\log\left(\frac{1}{\alpha\delta}\right)}{\epsilon}$. Optimizing the bound we set $\epsilon = \sqrt{\frac{h\log\left(\frac{1}{\alpha\delta}\right)}{OPT_{\mathcal{G}_\alpha}}}$ and get a revenue of

$$OPT_{\mathcal{G}_\alpha} - 8\sqrt{h\, OPT_{\mathcal{G}_\alpha} \;\; \log\left(\frac{1}{\alpha\delta}\right)},$$

which completes the proof. □

5 Attribute Auctions

We now consider applying our general bounds (Section 3) to attribute auctions. For attribute auctions an offer is a function from the publicly observable attribute of an agent to a take-it-or-leave-it price. As such, we identify such an offer with its *pricing function*. We begin by instantiating the results in Section 3 for market pricing auctions, in which we consider pricing functions that partition the attribute space into market segments and offer a fixed price in each. We show how one can use standard combinatorial dimensions in learning theory, e.g. the Vapnik-Chervonenkis (VC) dimension [3,11,16,30,36], in order to bound the complexity of these classes of offers. We then give an analysis for very general offer classes induced by general pricing functions over the attribute space that uses the notion of covers defined in Section 3.3.3.

5.1 Market Pricing

For attribute auctions, one natural class of pricing functions are those that segment bidders into *markets* in some simple way and then offer a single sale price in each market segment. For example, suppose we define $\mathcal{P}_k$ to be the set of functions that choose k bidders $b_1, \ldots, b_k$; use these as cluster centers to partition S into k markets based on distance to the nearest center in attribute space; and then offer a single price in each market. In that case, if we discretize prices to powers of $(1+\epsilon)$, then clearly the number of functions in the offer class $\mathcal{G}_k$ induced by the pricing class $\mathcal{P}_k$, is at most $n^k(\log_{1+\epsilon} h)^k$, so Corollary 2 implies that so long as $n \geq \frac{18h}{\epsilon^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln n + k\ln\left(\log_{1+\epsilon} h\right)\right]$ and assuming we can solve the optimization problem, then with probability at least $1-\delta$, we can get profit at least $(1-\epsilon)\,\mathrm{OPT}_{\mathcal{G}_k}$.

We can also consider more general ways of defining markets. Let C be any class of subsets of $\mathcal{X}$, which we will call *feasible markets*. For k a positive integer, we consider $F_{k+1}(C)$ to be the set of all pricing functions of the following form: pick k disjoint subsets $\mathcal{X}_1,\ldots,\mathcal{X}_k \subseteq \mathcal{X}$ from C, and $k+1$ prices $p_0,\ldots,p_k$ discretized to powers of $1+\epsilon$. Assign price p_i to bidders in $\mathcal{X}_i$, and price p_0 to bidders not in any of $\mathcal{X}_1,\ldots,\mathcal{X}_k$. For example, if $\mathcal{X}=\mathbb{R}^d$ a natural C might be the set of axis-parallel rectangles in $\mathbb{R}^d$. The specific case of $d=1$ was studied in [9]. One can envision more complex partitions, using the membership of a bidder in $\mathcal{X}_i$ as a basic predicate, and constructing any function over it (e.g., a decision list).

We can apply the results in Section 3 by using the machinery of VC-dimension to count the number of distinct such functions over any given set of bidders S. In particular, let $D=\mathrm{VCdim}(C)$ be the VC-dimension of C and assume $D<\infty$. Define $C[S]$ to be the number of distinct subsets of S induced by C. Then, from Sauer's Lemma $C[S] \leq \left(\frac{en}{D}\right)^D$, and therefore the number of different pricing functions in $F_k(C)$ over S is at most $\left(\log_{1+\epsilon} h\right)^k \left(\frac{en}{D}\right)^{kD}$. Thus applying Corollary 2 here we get:

Corollary 15 *Given a β-approximation algorithm $\mathcal{A}$ for optimizing over the offer class $\mathcal{G}_k$ induced by the class of pricing functions $F_k(C)$, then so long as $\mathrm{OPT}_{\mathcal{G}_k} \geq n$ and the number of bidders n satisfies*

$$n \geq \frac{18h\beta}{\epsilon^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon}\ln h\right) + kD\ln\left(\frac{ne}{D}\right)\right],$$

then with probability at least $1-\delta$, the profit of $RSO_{\mathcal{G}_k,\mathcal{A}}$ is at least $(1-\epsilon)\frac{\mathrm{OPT}_{\mathcal{G}_k}}{\beta}$.

The above lemma has "n" on both sides of the inequality. Simple algebra yields:

Corollary 16 *Given a β-approximation algorithm $\mathcal{A}$ for optimizing over the offer class $\mathcal{G}_k$ induced by the class of pricing functions $F_k(C)$, then so long as $\mathrm{OPT}_{\mathcal{G}_k} \geq n$ and the number of bidders n satisfies*

$$n \geq \frac{36h\beta}{\epsilon^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon}\ln h\right) + kD\ln\left(\frac{36kh\beta}{\epsilon^2}\right)\right],$$

then with probability at least $1-\delta$, the profit of $RSO_{\mathcal{G}_k,\mathcal{A}}$ is at least $(1-\epsilon)\frac{\mathrm{OPT}_{\mathcal{G}_k}}{\beta}$.

Proof: Since $\ln a \leq ab - \ln b - 1$ for all $a,b>0$, we obtain: $\frac{18kDh\beta}{\epsilon^2}\ln n \leq \frac{n}{2} + \frac{18kDh\beta}{\epsilon^2}\ln\left(\frac{36kDh\beta}{e\epsilon^2}\right)$. Therefore, it suffices to have: $n \geq \frac{n}{2} + \frac{18h\beta}{\epsilon^2}\Big[\ln\left(\frac{2}{\delta}\right) +$

$k \ln\left(\frac{1}{\epsilon}\ln h\right) + kD \ln\left(\frac{36kh\beta}{\epsilon^2}\right)\Big]$, so $n \geq \frac{36h\beta}{\epsilon^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon}\ln h\right) + kD\ln\left(\frac{36kh\beta}{\epsilon^2}\right)\right]$ suffices. □

For certain classes C we can get better bounds. In the following, denote by C_k the concept class of unions of at most k sets from C, and let L be $\lceil \log_{1+\epsilon} h \rceil$. If C is the class of intervals on the line, then the VC-dimension of C_k is $2k$, and so the number of different pricing functions in $F_k(C)$ over S is at most $L^k\left(\frac{en}{2k}\right)^{2k}$; also, if C is the class of all axis parallel rectangles in d dimensions, then the VC-dimension of C_k is $O(kd)$ [20]. In these cases we can remove the $\log k$ term in our bounds, which is nice because it means we can interpret our results (e.g., Corollary 16) as charging OPT a penalty for each market it creates. However, we do not know how to remove this $\log k$ term in general, since in general the VC-dimension of C_k can be as large as $2Dk\log(2Dk)$ (see [7,17]).

Corollary 16 gives a guarantee in the revenue of $\mathrm{RSO}_{\mathcal{G}_k,\mathcal{A}}$ so long as we have enough bidders. In the following, for $k \geq 0$ let $\mathrm{OPT}_k = \mathrm{OPT}_{\mathcal{G}_k}$. We can also use Corollaries 5 and 16 to show a bound that holds for all n, but with an additive loss term.

Theorem 17 *For any given value of n, k, ϵ, and δ, with probability at least $1-\delta$, the revenue of $RSO_{\mathcal{G}_k,\mathcal{A}}$ is*

$$\tfrac{1}{\beta}\left[(1-\epsilon)\,\mathrm{OPT}_k - h \cdot r_F(k, D, h, \epsilon, \delta)\right],$$

where $r_F(k, D, h, \epsilon, \delta) = O\left(\frac{kD}{\epsilon^2}\ln\left(\frac{kDh}{\epsilon\delta}\right)\right)$.

Proof: For simplicity, we show the proof for $\beta = 1$, the general case is similar. We prove the bound with the "$(1-\epsilon)$" term replaced by the term $\min\left(\frac{(1-\epsilon')^2}{1+\epsilon'}, 1-2\epsilon'\right)$, which then implies our desired result by simply using $\epsilon' = \frac{\epsilon}{3}$. If $n \geq \frac{36h}{\epsilon'^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon'}\ln h\right) + kD\ln\left(\frac{36kh}{\epsilon'^2}\right)\right]$, then the desired statement follows directly from Corollary 16. Otherwise, consider first the case when we have $\mathrm{OPT}_k \geq \frac{4h}{\epsilon'^2(1-\epsilon')}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon'}\ln h\right) + kD\ln\left(\frac{ne}{D}\right)\right]$. Let g_i be the optimal offer in $\mathcal{G}_k$ over S_i, for $i = 1, 2$, and let g_{OPT} be the optimal offer in $\mathcal{G}_k$ over S (and so $g_i(S_i) \geq g_{\mathrm{OPT}}(S_i)$). From Corollary 5, we have $g_{\mathrm{OPT}}(S_i) \geq \frac{2h}{\epsilon'^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon'}\ln h\right) + kD\ln\left(\frac{ne}{D}\right)\right]$, for $i = 1, 2$. So, $g_i(S_i) \geq \frac{2h}{\epsilon'^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon'}\ln h\right) + kD\ln\left(\frac{ne}{D}\right)\right]$. Using again Corollary 5, we obtain $g_i(S_j) \geq \frac{1-\epsilon'}{1+\epsilon'}g_i(S_i)$ for $j \neq i$, which then implies the desired result. To complete the proof notice that if both $\mathrm{OPT}_k \leq \frac{4h}{\epsilon'^2(1-\epsilon')}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{1}{\epsilon'}\ln h\right) + kD\ln\left(\frac{ne}{D}\right)\right]$ and $n \leq \frac{4h}{\epsilon'^2}\left[\ln\left(\frac{2}{\delta}\right) + k\ln\left(\frac{2}{\epsilon'}\ln h\right) + kD\ln\left(\frac{4kh}{\epsilon'^2}\right)\right]$, then we easily get the desired statement. □

Finally, as in Theorem 6 we can extend our results to use structural risk minimization, where we want the algorithm to optimize over k, by viewing the additive loss term, $h \cdot r_F(\cdot)$, as a penalty function.

Theorem 18 *Let $\bar{\mathcal{G}}$ be the sequence $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n$ of offer classes induced by the sequence of classes of pricing functions $F_1(C), F_2(C), \dots, F_n(C)$. Then for any value of n, ϵ and δ with probability $1-\delta$ the revenue of RSO-SRM$_{\bar{\mathcal{G}},\mathsf{pen}}$ is*

$$\max_k ((1-\epsilon)\,\mathrm{OPT}_k - h \cdot r_F(k, D, h, \epsilon, \delta)),$$

where $\mathsf{pen}(F_k(C)) = \frac{h}{2} \cdot r_F(k, D, h, \epsilon, \delta) = O\left(\frac{kD}{\epsilon^2} \ln\left(\frac{kDh}{\epsilon\delta}\right)\right)$.

To illustrate the tightness of Theorem 17, notice that even for the special case of pricing using interval functions (the case of $d = 1$ studied in [9]), the following lower bound holds.

Theorem 19 *Let $\mathcal{X} = \mathbb{R}$ and let C_k be the class of k intervals over $\mathcal{X}$. Then there is no incentive compatible mechanism whose expected revenue is at least $\frac{3}{4}\,\mathrm{OPT}_k - o(kh)$.*

That is, an additive loss linear in kh is necessary in order to achieve a multiplicative ratio of at least 3/4.

Proof: Consider $\frac{kh}{2}$ bidders with distinct attributes (for instance, say bidder i has attribute i), each of whom independently has a $\frac{1}{h}$ probability of having valuation h and a $1-\frac{1}{h}$ probability of having valuation 1. Then, any incentive-compatible mechanism has expected profit at most $\frac{kh}{2}$ because for any given bidder and any given proposed price, the expected profit (over randomization in the bidder's valuation) is at most 1. However, there is at least a 50% chance we will have at least $\frac{k}{2}$ bidders of valuation h, and in that case OPT_k can give $\frac{k}{2} - 1$ of those bidders a price of h and the rest a price of 1 for an expected profit of $\left(\frac{k}{2} - 1\right) h + \left(\frac{kh}{2} - \frac{k}{2} + 1\right) 1 = kh - h - \frac{k}{2} + 1$. On the other hand even if that does not occur, we always have $\mathrm{OPT}_k \geq \frac{kh}{2}$. So, the expected profit of OPT_k is at least $3\frac{kh}{4} - \frac{h}{2} - \frac{k}{4}$. Thus, the profit of the incentive-compatible mechanism is at most $\frac{3}{4}\,\mathrm{OPT}_k - \frac{kh}{16} + o(kh)$. □

We note that a similar lower bound holds for most base classes. Also for the case of intervals on the line, both our auction and the auction in [9] match this lower bound up to constant factors.

5.2 General Pricing Functions over the Attribute Space

In this section we generalize the results in Section 5.1 in two ways: we consider general classes of pricing functions (not just piecewise-constant functions defined over markets), and we remove the need to discretize by instead using the covering arguments discussed in Section 3.3.3. This allows us to consider offers based on linear or quadratic functions of the attributes, or perhaps functions that divide the attribute space into markets and use pricing functions are linear in the attributes (rather than constant) in each market. The key point of this section is that we can bound the size of the L_1 multiplicative cover in an attribute auction in terms of natural quantities.

Assume in the following that $\mathcal{X} \subseteq \mathbb{R}^d$, let $\mathcal{P}$ be a fixed class of pricing functions over the attribute space $\mathcal{X}$ and let $\mathcal{G}$ be the induced class of offers. Let $\mathcal{P}_d$ be the class of decision surfaces (in $\mathbb{R}^{d+1}$) induced by $\mathcal{P}$: that is, to each $q \in \mathcal{P}$ we associate the set of all $(x, v) \in \mathcal{X} \times [1, h]$ such that $q(x) \leq v$. Also, let us denote by D the VC-dimension of class $\mathcal{P}_d$. We can then show that:

Theorem 20 *Given the offer class $\mathcal{G}$ and a β-approximation algorithm $\mathcal{A}$ for optimizing over $\mathcal{G}$, then so long as $\mathrm{OPT}_{\mathcal{G}} \geq n$ and the number of bidders n satisfies*

$$n \geq \frac{154h\beta}{\epsilon^2}\left[\ln\left(\frac{2}{\delta}\right) + D\ln\left(\frac{154h\beta}{\epsilon^2}\left(\frac{12}{\epsilon}\ln h + 1\right)\right)\right],$$

then with probability at least $1-\delta$, the profit of $RSO_{(\mathcal{G},\mathcal{A})}$ is at least $(1-\epsilon)\frac{\mathrm{OPT}_{\mathcal{G}}}{\beta}$.

The key to the proof is to exhibit an L_1 multiplicative cover of $\mathcal{G}$ whose size is exponential in D only, and then to apply Corollary 10.

Proof: Let $\alpha = \frac{\epsilon}{12}$. For each bidder (x, v) we conceptually introduce $O(\frac{1}{\alpha}\ln h)$ "phantom bidders" having the same attribute value x and bid values $1, (1+\alpha), (1+\alpha)^2, \cdots, h$. Let S^* be the set S together with the set of all phantom bidders; let $n^* = |S^*|$. Let Split be the set of possible splittings of S^* with surfaces from $\mathcal{P}_d$. We clearly have $|\mathsf{Split}| \leq \mathcal{P}_d[S^*]$. For each element $s \in \mathsf{Split}$ consider a representative function in $\mathcal{G}$ that induces splitting s in terms of its winning bidders, and let $\mathsf{Split}_{\mathcal{G}}$ be the set of these representative functions. Let $\mathcal{G}'$ be the offer class induced by the pricing class $\mathsf{Split}_{\mathcal{G}}$. Notice that $\mathcal{G}'$ is actually an L_1 multiplicative α-cover for $\mathcal{G}$ with respect to S, since for every offer in $\mathcal{G}$ there is a offer in $\mathcal{G}'$ that extracts nearly the same profit from every bidder; i.e., for every offer in $g \in \mathcal{G}$, there exists $g' \in \mathcal{G}'$ such that for every $(x, v) \in S$, we have both $g'((x, v)) \leq (1+\alpha)g((x, v))$ and $g((x, v)) \leq (1+\alpha)g'((x, v))$. From Sauer's lemma we know $|\mathsf{Split}_{\mathcal{G}}| \leq \left(\frac{n^* e}{D}\right)^D$, and applying Corollary 10, we finally get the desired statement by using simple algebra as in Corollary 16. □

The above theorem is the analog of Corollary 2. Using it and Theorem 9, it is easy to derive a bound that holds for all n (i.e., the analog of Theorem 17). One can further easily extend these results to get bounds for the corresponding SRM auction (as done in Theorem 18).

There has been relatively little work on the algorithmic question of computing optimal pricing functions in general attribute spaces. However, for single-dimentional attributes and piece-wise constant pricing functions [9] discusses an optimal polynomial time dynamic program. For single-dimentional attributes and monotone pricing functions, [2] gives a polynomial time dynamic program. The problem of computing the optimal of linear pricing function over m-dimentional attributes generalizes the problem of item-pricing (m distinct items) for single-minded combinatorial consumers (see Section 6.4) that has been shown to be hard to approximate to better than a $\log^{\delta}(m)$ factor for some $\delta > 0$ [15].

6 Combinatorial Auctions

Combinatorial auctions have received much attention in recent years because of the difficulty of merging the algorithmic issue of computing an optimal outcome with the game-theoretic issue of incentive compatibility. To date, the focus primarily has been on the problem of optimizing social welfare: partitioning a limited supply of items among bidders to maximize the sum of their valuations. We consider instead the goal of profit maximization for the seller in the case that the items for sale are available in unlimited supply. [7] We consider the general version of the combinatorial auction problem as well as the special cases of *unit-demand* bidders (each bidder desires only singleton bundles) and *single-minded* bidders (each bidder has a single desired bundle).

It is interesting to restrict our attention to the case of item-pricing, where the auctioneer intuitively is attempting to set a price for each of the distinct items and bidders then choose their favorite bundle given these prices. Item-pricing is without loss of generality for the unit-demand case, and general bundle-pricing can be realized with an auction with $m' = 2^m$ "items", one for each of possible bundle of the original m items. [8]

First notice that if the set of allowable item pricings are constrained to be integral, $\mathcal{G}_{\mathbb{Z}}$, then clearly there are at most $|\mathcal{G}_{\mathbb{Z}}| = (h+1)^m$ possible item

[7] Other work focusing on profit maximization in combinatorial auctions include Goldberg and Hartline [21], Hartline and Koltun [29], Guruswami et al. [28], Likhodedov and Sandholm [32], and Balcan et al. [6].

[8] We make the assumption that all desired bundles contain at most one of each item. This assumption can be easily relaxed and our results apply given any bound on the number of copies of each item that are desired by any one consumer. Of course, this reduction produces an exponential blowup in the number of items.

pricings. By Corollary 2 we get that $\tilde{O}\left(\frac{hm}{\epsilon^2}\right)$ bidders are sufficient to achieve profit close to $\mathrm{OPT}_{\mathcal{G}_{\mathbb{Z}}}$. Generally it is possible to do much better if non-integral item-pricings are allowed, i.e., $\mathrm{OPT}_{\mathcal{G}}(S) \gg \mathrm{OPT}_{\mathcal{G}_{\mathbb{Z}}}(S)$. In these settings we can still get good bounds following the guidelines established in Section 3.3, by either considering an offer class $\mathcal{G}'$ induced by discretization (see Section 6.1), or from counting possible outcomes in $\mathcal{G}_{\mathcal{A}}$ (see Section 6.2). A summary of our results is given in Table 1.

	general	unit-demand	single-minded
$\lvert\mathcal{G}'\rvert$	$O(\log^m_{1+\epsilon^2} \frac{nm}{\epsilon})$	$O(\log^m_{1+\epsilon^2} \frac{n}{\epsilon})$	$O(\log^m_{1+\epsilon} \frac{nm}{\epsilon})$
$\lvert\mathcal{G}_{\mathcal{A}}\rvert$	$n^m 2^{2m^2}$	$n^m(m+1)^{2m}$	$(n+m)^m$

Table 1
Size of offer classes for combinatorial auctions.

We can apply Theorem 1 and Corollary 2 to the sizes of the offer classes in Table 1 to get bounds on the profit of random sampling auctions for combinatorial item pricing. In particular, using Corollary 2 we get that $\tilde{O}\left(\frac{hm^2}{\epsilon^2}\right)$ bidders are sufficient to achieve revenue close to the optimum item-pricing in the general case, and $\tilde{O}\left(\frac{hm}{\epsilon^2}\right)$ bidders are sufficient for the unit-demand case. Also, by using Theorem 1 instead of Corollary 2 we can replace the condition on the number of bidders with a condition on $\mathrm{OPT}_{\mathcal{G}}$, which gives a factor of m improvement on the bound given by [21].

As before we let $h = \max_{g \in \mathcal{G}, i \in S} g(i)$. In particular, this implies that $\mathrm{OPT}_{\mathcal{G}} \geq h$ which will be important later in this section.

6.1 Bounds via Discretization

As shown in Section 3.3.1, we can obtain good bounds if we are willing to optimize over a set $\mathcal{G}'$ of offers induced by a small set of discretized prices satisfying that $\mathrm{OPT}_{\mathcal{G}'}$ is close to $\mathrm{OPT}_{\mathcal{G}}$. Prior to this work, [29] shows how to construct discretized classes $\mathcal{G}'$ with $\mathrm{OPT}_{\mathcal{G}'} \geq \frac{1}{1+\epsilon}\,\mathrm{OPT}_{\mathcal{G}}$ and size $O(m^m \log^m_{1+\epsilon} \frac{n}{\epsilon})$ for the unit-demand case and size $O(\log^m_{1+\epsilon} \frac{nm}{\epsilon})$ for the single-minded case. Nisan [34] gives the basic argument necessary to generalize these results to obtain the result in Theorem 21 which applies to combinatorial auctions in general. We note in passing that Theorem 21 allows for generalization and improvement of the computational results of [29]. The discretization results we obtain are summarized in the first row of Table 1.

Let $\mathbf{p} = (p_1, \ldots, p_m)$ be an item-pricing of the m items. Let $g_{\mathbf{p}}$ correspond to the offering pricing $\mathbf{p}$. The following is the main result of this section.

Theorem 21 *Let k be the size of the maximum desired bundle. Let $\mathbf{p}'$ be the optimal discretized price vector that uses item prices equal to* 0 *or powers of*

$(1+\epsilon)$ in the range $\left[\frac{h\epsilon}{nk}, h\right]$ and let $\mathbf{p}^*$ be the optimal price vector. Then we have:

$$g_{\mathbf{p}'}(S) \geq (1-2\sqrt{\epsilon})g_{\mathbf{p}^*}(S).$$

Proof: Let $\delta = \sqrt{\epsilon}$. For the optimal price vector $\mathbf{p}^*$ with item j priced at p_j^* (i.e., $g_{\mathbf{p}^*}(S) = \text{OPT}_{\mathcal{G}}$), consider a price vector $\mathbf{p}$ with p_j in $[(1-\delta)p_j^*, (1-\delta+\delta^2)p_j^*]$ if $p_j^* \geq \frac{h\delta^2}{nk}$ and 0 otherwise, where $p_j = (1+\epsilon)^k$ for some integer k (note that such a price vector always exists). We show now that $g_{\mathbf{p}}(S) \geq (1-2\sqrt{\epsilon})g_{\mathbf{p}^*}(S)$, which clearly implies the desired result.

Let J be a multi-set of items and $\text{Profit}(J) = \sum_{j\in J} p_j^*$ be the payment necessary to purchase bundle J under pricing $\mathbf{p}^*$. Define $R_j = p_j^* - p_j$. Thus we have:

$$(\delta-\delta^2)p_j^* \leq R_j \leq \max\{\delta p_j^*, \tfrac{\delta^2 h}{nk}\} \leq \delta p_j^* + \tfrac{\delta^2 h}{nk}.$$

This implies that for any multiset J with $|J| \leq k$, we have the following upper and lower bounds:

$$\sum_{j\in J} R_j \geq (\delta-\delta^2)\text{Profit}(J)\ , \tag{1}$$

$$\sum_{j\in J'} R_j \leq \delta\text{Profit}(J') + \tfrac{h\delta^2}{n}. \tag{2}$$

Let J_i^* and J_i be the bundles that bidder i prefers under pricing $\mathbf{p}^*$ and $\mathbf{p}$, respectively. Consider bidder i who switches from bundle J_i^* to bundle J_i when the item prices are decreased from $\mathbf{p}^*$ to $\mathbf{p}$. This implies that:

$$\sum_{j\in J_i^*} R_j \leq \sum_{j\in J_i} R_j.$$

Combining this with equations (1) and (2) and canceling a common factor of δ we see that:

$$(1-\delta)\text{Profit}(J_i^*) \leq \text{Profit}(J_i) + \tfrac{h\delta}{n}.$$

Summing over all bidders i, we see that the total profit under our new pricing $\mathbf{p}$ is at least $(1-\delta)\,\text{OPT}_{\mathcal{G}} - h\delta$. Since $\text{OPT}_{\mathcal{G}} \geq h$, we finally obtain that the profit under $\mathbf{p}$ is at least $(1-2\delta)\,\text{OPT}_{\mathcal{G}}$. □

Note that we can now apply Theorem 21 by letting $\mathcal{G}'$ be the offer class induced by the class of item prices equal to 0 or powers of $(1+\epsilon)$ in the range $\left[\frac{h\epsilon}{nk}, h\right]$ (where k bounds the maximum size of a bundle). Using Theorem 1 we obtain the following guarantee:

Corollary 22 *Given a β-approximation algorithm $\mathcal{A}$ optimizing over $\mathcal{G}'$, then with probability at least $1-\delta$, the profit of $RSO_{\mathcal{G}',\mathcal{A}}$ is at least $(1-3\epsilon)\text{OPT}_{\mathcal{G}}/\beta$*

so long as

$$\mathrm{OPT}_{\mathcal{G}'} \geq \tfrac{18h\beta}{\epsilon^2}\left(m\ln(\log_{1+\epsilon^2} nk) + \ln\left(\tfrac{2}{\delta}\right)\right).$$

6.2 Bounds via Counting

We now show how to use the technique of counting possible outcomes (See Section 3.3.2) to get a bound on the performance of the random sampling auction with an algorithm $\mathcal{A}$ for item-pricing. This approach calls for bounding $|\mathcal{G}_{\mathcal{A}}|$, the number of different pricing schemes $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ can possibly output. Our results for this approach are summarized in the second row of Table 1.

Recall that bidder i's utility for a bundle J given pricing $\mathbf{p}$ is $u_i(J,\mathbf{p}) = v_i(J) - \sum_{j\in J} p_j$. We now make the following claim about the regions of the space of possible pricings, $\mathbb{R}^m_+$, in which bidder i's most desired bundle is fixed.

Claim 2 *Let* $P_i(J) = \{\mathbf{p} \mid \forall J', u_i(J,\mathbf{p}) \geq u_i(J',\mathbf{p})\}$. *The set* $P_i(J,\mathbf{p})$ *is a polytope.*

Proof: This follows immediately from the observation that the region $P_i(J)$ is convex and the only way to pack convex regions into space is if they are polytopes.

To show that $P_i(J)$ is convex, suppose the allocation to a particular bidder for $\mathbf{p}$ and $\mathbf{p}'$ are the same, J. Then for any other bundle J' we have:

$$v_i(J) - \sum_{j\in J} p_j \geq v_i(J') - \sum_{j\in J'} p_j$$

and

$$v_i(J) - \sum_{j\in J} p'_j \geq v_i(J') - \sum_{j\in J'} p'_j.$$

If we now consider any price vector $\alpha\mathbf{p} + (1-\alpha)\mathbf{p}'$, for $\alpha \in [0,1]$, these imply:

$$v_i(J) - \sum_{j\in J}(\alpha p_j + (1-\alpha)p'_j) \geq v_i(J') - \sum_{j\in J'}(\alpha p_j + (1-\alpha)p'_j).$$

This clearly implies that this agent prefers allocation J on any convex combination of $\mathbf{p}$ and $\mathbf{p}'$. Hence the region of prices for which the agent prefers bundle J is convex. □

The above claim shows that we can divide the space of pricings into polytopes based on an agent's most desirable bundle. Consider fixing an outcome, i.e., the bundles $J_1, \ldots, J_n$, obtained by agents $1, \ldots, n$, respectively. This outcome occurs for pricings in the intersection $\bigcap_{i\in S} P_i(J_i)$.

Definition 2 *For a set of agents S, let Verts_S denote the set of vertices of the polytopes that partition the space of prices by the allocation produced. I.e., $\mathsf{Verts}_S = \{\mathbf{p}$ such that $\mathbf{p}$ is a vertex of the polytope containing $\bigcap_{i \in S'} P_i(J_i)$ for some $i \in S' \subset S$ and bundles $J_i\}$.*

Claim 3 *For $S' \subseteq S$ we have $\mathsf{Verts}_{S'} \subseteq \mathsf{Verts}_S$.*

Proof: Follows immediately from the definition of Verts_S and basic properties of polytopes. □

Now we consider optimal pricings. Note that when fixing an allocation $J_1, \ldots, J_n$ we are looking for an optimal price point within the polytope that gives this allocation. Our objective function for this optimization is linear. Let n_j be the number of copies of item j allocated by the allocation. The seller's payoff for prices $\mathbf{p} = (p_1, \ldots, p_m)$ is $\sum_j p_j n_j$. Thus, all optimal pricings of this allocation lie on facets of the polytope and in particular there is an optimal pricing that is at a vertex of the polytope. Over the space of all possible allocations, all optimal pricings are on facets of the allocation defining polytopes and there exists an optimal pricing that is at a vertex of one of the polytopes.

Lemma 23 *Given an algorithm $\mathcal{A}$ that always outputs a vertex of the polytope then $\mathcal{G}_\mathcal{A} \subseteq \mathsf{Verts}_S$.*

Proof: This follows from the fact that $\mathrm{RSO}_{(\mathcal{G},\mathcal{A})}$ runs $\mathcal{A}$ on a subset S' of S which has $\mathsf{Verts}_{S'} \subseteq \mathsf{Verts}_S$. $\mathcal{A}$ must pick a price vector from $\mathsf{Verts}_{S'}$. By Claim 3 this price vector must also be in Verts_S. This gives the lemma. □

We now discuss getting a bound on Verts_S for n agents, m distinct items, and various types of preferences.

Theorem 24 *We have the following upper bounds on $|\mathsf{Verts}_S|$:*

(1) $(n+m)^m$ for single-minded preferences.
(2) $n^m(m+1)^{2m}$ for unit-demand preferences.
(3) $n^m 2^{2m^2}$ for arbitrary preferences.

Proof: We consider how many possible bundles, M, an agent might obtain as a function of the pricing. An agent with single-minded preferences will always obtain one of $M_s = 2$ bundles: either their desired bundle or nothing (the empty bundle). An agent with unit-demand preferences receives one of the m items or nothing for a total of $M_u = m + 1$ possible bundles. An agent with general preferences receives one of the $M_g = 2^m$ possible bundles.[9]

[9] Here we make the assumption that desired bundles are simple sets. If they are actually multi-sets with bounded multiplicity k, then the agent could receive one of at most $M_g = (k+1)^m$ bundles.

We now bound the number of hyperplanes necessary to partition the pricing space into M convex regions (e.g., that specify which bundle the agent receives). For convex regions, each pair of regions can meet in at most one hyperplane. Thus, the total number of hyperplanes necessary to partition the pricing space into regions is at most $\binom{M}{2}$. Of course we wish to restrict our pricings to be non-negative, so we must add m additional hyperplanes at $p_j = 0$ for all j.

For all n agents, we simply intersect the regions of all agents. This does not add any new hyperplanes. Furthermore, we only need to count the m hyperplanes that restrict to non-negative pricings once. Thus, the total number of hyperplanes necessary for specifying the regions of allocation for n agents with M convex regions each, is $K = n\binom{M}{2} + m$. Thus, $K_s = n + m$, $K_u \leq n\binom{m+1}{2} + m \leq n(m+1)^2$, and $K_g \leq n\binom{2^m}{2} + m \leq n2^{2m}$ (for $m \geq 2$).

Of course, K hyperplanes in m dimensional space intersect in at most $\binom{K}{m} \leq K^m$ vertices. Not all of these intersections are vertices of polytopes defining our allocation, still K^m is an upper bound on the size of Verts_S. Plugging this in gives us the desired bounds of $(n+m)^m$, $n^m(m+1)^{2m}$, and $n^m 2^{2m^2}$ respectively for single-minded, unit-demand, and general preferences. □

We note that are above arguments apply to approximation algorithms that always output a price corresponding to the vertex of a polytope as well. Though we do not consider this direction here, it is entirely possible that it is not computationally difficult to post-process the solution of an algorithm that is not a vertex of a polytope to get a solution that is on a vertex of a polytope. [10] This would further motivate the analysis above. If for some reason, restricting to algorithms that return vertices is undesirable, it is possible to use cover arguments on the set of vertices we obtain when we add additional hyperplanes corresponding to the discretization of the preceding section.

6.3 Combinatorial Auctions: Lower Bounds

We show in the following an interesting lower bound for combinatorial auctions. [11] Notice that our upper bounds and this lower bound are quite close.

Theorem 25 *Fix m and h. There exists a probability distribution on unit-demand single-minded agents such that the expected revenue of any incentive*

[10] Notice that this is not immediate because of the complexity of representing an agent's combinatorial valuation.

[11] This proof follows the standard approach for lower bounds for revenue maximizing auctions that was first given by Goldberg et al. in [24].

compatible mechanism is at most $\frac{mh}{2}$ *whereas the expected revenue of* OPT *is at least* $0.7mh$.

Thus, this theorem states that in order to achieve a close multiplicative ratio with respect to OPT, one must have additive loss $\Omega(mh)$.

Proof: Consider the following probability distribution over valuations of agents preferences. Assume we have $n = \frac{mh}{2}$ agents in total, and $\frac{h}{2}$ agents desire item j only, $j \in \{1, \cdots m\}$.[12] Each of these agents has valuation h with probability $\frac{1}{h}$ and valuation 1 with probability $1 - \frac{1}{h}$.

Notice now any incentive-compatible mechanism has expected profit at most n. To see this, note that for each bidder, any proposed price has expected profit (over the randomization in the selection of his valuation) of at most 1. Moreover, the expected profit of $\mathrm{OPT}_{\mathcal{G}}$ is at least $n + \frac{mh}{8}$. For each item j, there is a $1-(1-\frac{1}{h})^{h/2} \approx 0.4$ probability that some bidder has valuation h. For those items, $\mathrm{OPT}_{\mathcal{G}}$ gets at least a profit of h. For the rest, $\mathrm{OPT}_{\mathcal{G}}$ gets a profit of $\frac{h}{2}$. So, overall, $\mathrm{OPT}_{\mathcal{G}}$ gets an expected profit of at least $0.4mh + 0.6m(h/2) = 0.7h$. All these together imply the desired result. □

6.4 Algorithms for Item-pricing

Given standard complexity assumptions, most item-pricing problems are not polynomial time solvable, even for simple special cases. We review these results here. We restrict our attention to the unlimited supply special case, though some of the work we mention also considers limited supply item-pricing. Algorithmic pricing problems in this form were first posed by Guruswami et al. [28] though item-pricing for unit-demand consumers with several alternative payment rules (i.e., rules that do not represent quasi-linear utility maximization) were independently considered by Aggarwal et al. [1].

For consumers with single-minded preferences, [28] gives a simple $O(\log mn)$ approximation algorithm. Demaine et al. [15] show the problem to be hard to approximate to better than a $\log^{\delta}(m)$ factor for some $\delta > 0$. Both Briest and Krysta [14] and Grigoriev et al. [27] proved that optimal pricing is weakly NP-hard for the special case known as "the highway problem" where there is a linear order on the items and all desired bundles are for sets of consecutive items (actually this hardness result follows for the more specific case where the desired bundles for any two agents, S_i and $S_{i'}$, satisfy one of: $S_i \subseteq S_{i'}$, $S_{i'} \subseteq S_i$, or $S_i \cup S_{i'} = \emptyset$). In the case when the cardinality of the desired bundles are bounded by k, Briest and Krysta [14] give an $O(k^2)$ approximation algorithm, which is improved to $O(k)$ by Balcan and Blum [5]. Finally, when the number

[12] Notice that these preferences are both unit-demand and single-minded.

of distinct items for sale, m, is constant, Hartline and Koltun [29] show that it is possible to improve on the trivial $O(n^m)$ algorithm by giving a near-linear time approximation scheme. Their approximation algorithm is actually an exact algorithm for the problem of optimizing over a discretized set of item prices $\mathcal{G}'$ which is directly applicable to our auction $\mathrm{RSO}_{(\mathcal{G}',\mathcal{A})}$, discussed above.

For consumers with unit-demand preferences, [28] (and [1] essentially) give a trivial logarithmic approximation algorithm and show that the optimization problem is APX-hard (meaning that standard complexity assumptions imply that there does not exist a polynomial time approximation scheme (PTAS) for the problem). Again, Hartline and Koltun [29] show how to improve on the trivial $O(n^m)$ algorithm in the case where the number of distinct items for sale, m, is constant. They give a near-linear time approximation scheme that is based on considering a discretized set of item prices; however, the discretization of Nisan [34] that we discussed above gives a significant improvement on their algorithm and also generalizes it to be applicable to the problem of item-pricing for consumers with general combinatorial preferences.

7 Conclusions, Discussion, and Open Problems

In this work we have made an explicit connection between machine learning and mechanism design. In doing so, we obtain a *unified* approach to considering a variety of profit maximizing mechanism design problems including many that have been previously considered in the literature.

Some of our techniques give suggestions for the *design* of mechanisms and others for their *analysis*. In terms of design, these include the use of discretization to produce smaller function classes, and the use of structural-risk-minimization to choose an appropriate level of complexity of the mechanism for a given set of bidders. In terms of analysis, these include both the use of basic sample-complexity arguments, and the notion of multiplicative covers for better bounding the true complexity of a given class of offers.

Our results substantially generalize the previous work on random sampling mechanisms by both broadening the applicability of such mechanisms and by simplifying the analysis. Our bounds on random sampling auctions for digital goods not only show how the auction profit approaches the optimal profit, but also weaken the required assumptions of [26] by a constant factor. Similarly, for random sampling auctions for multiple digital goods, our unified analysis gives a bound that weakens the assumptions of [21] by a factor of more than m, the number of distinct items. This multiple digital good auction problem is a special case of the a more general unlimited supply combinatorial auction problem for which we obtain the first positive worst-case results by showing that it is possible to approximate the optimal profit with an incentive-compatible mechanism. Furthermore, unlike the case for combinatorial auctions for social welfare maximization, our incentive-compatible mechanisms can be easily based on approximation algorithms instead of exact ones.

We have also explored the attribute auction problem that was proposed in [9] for 1-dimensional attributes in a much more general setting: the attribute values can be multi-dimensional and the target pricing functions considered can be arbitrarily complex. We bound the performance of random sampling auctions as a function of the complexity of the target pricing functions.

Our random sampling auctions assume the existence of exact or approximate pricing algorithms. Solutions to these pricing problem have been proposed for several of our settings. In particular, optimal item-pricings for combinatorial auctions in the single-minded and unit-demand special cases have been considered in [5,14,29,28]. On the other hand for attribute auctions, many of the clustering and market-segmenting pricing algorithms have yet to be considered at all.

Open Problems: Probably the most important direction for future work is in relaxing the assumption that the items for sale are available in unlimited supply. In the random sampling framework, we propose the following mechanism: randomly partition the bidders into two sets, evenly divide the supply among the two sets, compute the optimal *envy-free*[13] offer for the two partitions, and apply the offer to the opposite partition. Of course, an offer g that is envy-free for S_1 may not necessarily be envy-free for S_2. There are several approaches that may work here. First, we could artificially deplete the supply by a constant factor and ask for an offer that is envy-free for the depleted supply. Then it may be possible to argue that it is envy-free for both S_1 and S_2 with high probability. Another option would be to take the bidders of S_2 in an arbitrary (or random) order and allow them to take their preferred outcome suggested by the offer constrained such that their preference is feasible given the remaining supply. It is easy to see that the technique outlined above results in an incentive compatible mechanism. Is it also close to optimal? Borgs et al. have successfully applied this latter approach to limited supply multi-unit auctions for bidders with budgets [12].

It is possible to further generalize the feasibility constraints imposed by limited supply to arrive at the general single-parameter agent auction problem (See e.g., [23] for a precise definition). This abstract problem can be viewed as

[13] To generalize envy-freedom [28] to attribute auctions, declare an offer $g \in \mathcal{G}$ envy-free for bidders S if there is enough supply such that all bidders that have strictly positive utility for their preferred outcome under g can simultaneously be satisfied without creating an infeasible outcome.

auctioning a service to a number of agents where the service provider must pay a cost that is a function of the agents served. In its full generality, this cost function could be arbitrary. The possibly asymmetric cost function can be viewed as endowing the agents with public attributes, or the agents could have additional attributes. A very interesting direction for future research is in determining for what classes of cost functions the general problem of profit maximization in this setting can be solved.

The final direction of investigation we propose is that of generalizing the special purpose bounds we obtain for digital good auctions (Section 4) to our general unlimited supply setting (Section 3). Recall that in for digital goods and indistinguishable bidders we were able to employ a telescoping argument to reduce the additive loss term to $O(h)$ which is optimal up to a constant factor. This takes advantage of the properties of take-it-or-leave-it prices: that the payoff for any given bidder is upper-bounded by the offer price. This allows us to use non-uniform bounds on the payoffs of the different pricing functions and these non-uniform bounds telescope. Can some form of this telescoping be generalized to attribute auctions, combinatorial auctions, or our general bounds? It would be also interesting to see if one can use some of the very recent techniques and ideas used in the context of learning theory and empirical processes (see e.g. [13,8,31]) to get better bounds for our mechanism design setting. In particular, it would be interesting to investigate data dependent bounding techniques in this setting.

References

[1] G. Aggarwal, T. Feder, R. Motwani, and A. Zhu. Algorithms for multi-product pricing. In *Proceedings of the International Colloquium on Automata, Languages, and Programming*, pages 72–83, 2004.

[2] G. Aggarwal and J. Hartline. Knapsack Auctions. In *Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms*, 2006.

[3] M. Anthony and P. Bartlett. *Neural Network Learning: Theoretical Foundations*. Cambridge University Press, 1999.

[4] B. Awerbuch, Y. Azar, and A. Meyerson. Reducing truth-telling online mechanisms to online optimization. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*, 2003.

[5] M.-F. Balcan and A. Blum. Approximation Algorithms and Online Mechanisms for Item Pricing. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, 2006. An earlier version available as Technical Report, CMU-CS-05-176.

[6] M.-F. Balcan, A. Blum, and Y. Mansour. Single price mechanisms for revenue maximization in unlimited supply combinatorial auctions. Technical Report CMU-CS-07-111, February 2007.

[7] P. Bartlett and W. Maass. Vapnik Chervonenkis Dimension of Neural Nets. In *The Handbook of Brain Theory and Neural Networks*. MIT Press, 2003.

[8] P. Bartlett and S. Mendelson. Rademacher and Gaussian Complexities Risk Bounds and Structural Results. *Journal of Machine Learning Research*, 54(3):463–482, 2002.

[9] A. Blum and J. Hartline. Near-Optimal Online Auctions. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms*, pages 1156 – 1163, 2005.

[10] A. Blum, V. Kumar, A. Rudra, and F. Wu. Online Learning in Online Auctions. In *Proceedings of the 14th ACM-SIAM Symposium on Discrete Algorithms*, pages 137 – 146, 2003.

[11] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth. Learnability and the Vapnik-Chervonenkis Dimension. *Journal of the ACM*, 36:929– 965, 1989.

[12] C. Borgs, J. T. Chayes, N. Immorlica, M. Mahdian, and A. Saberi. Multi-unit auctions with budget-constrained bidders. In *Proceedings of the 6th ACM Conference on Electronic Commerce*, pages 44–51, 2005.

[13] O. Bousquet, S. Boucheron, and G. Lugosi. Theory of Classification: A Survey of Recent Advances. *ESAIM: Probability and Statistics*, 2005.

[14] P. Briest and P. Krysta. Single-Minded Unlimited Supply Pricing on Sparse Instances. In *Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms*, 2006.

[15] E. Demaine, U. Feige, M.T. Hajiaghayi, and M. Salavatipour. Combination Can Be Hard: Approximability of the Unique Coverage Problem . In *Proceedings of the 17th ACM-SIAM Symposium on Discrete Algorithms*, 2006.

[16] L. Devroye, L. Gyorfi, and G. Lugosi. *A Probabilistic Theory of Pattern Recognition*. Springer-Verlag, 1996.

[17] L. Devroye and G. Lugosi. *Combinatorial Methods in Density Estimation*. Springer-Verlag, 2001.

[18] U. Feige, A. Flaxman, J. Hartline, and R. Kleinberg. On the Competitive Ratio of the Random Sampling Auction. In *Proc. 1st Workshop on Internet and Network Economics*, 2005.

[19] A. Fiat, A. Goldberg, J. Hartline, and A. Karlin. Competitive Generalized Auctions. In *Proceedings 34th ACM Symposium on the Theory of Computing*, pages 72 – 81, 2002.

[20] P. Fischer and S. Kwek. Minimizing Disagreement for Geometric Regions Using Dynamic Programming, with Applications to Machine Learning and Computer Graphics. Technical Report eC-TR-96-004, 1996.

[21] A. Goldberg and J. Hartline. Competitive Auctions for Multiple Digital Goods. In *Proceedings of the 9th Annual European Symposium on Algorithms*, pages 416 – 427, 2001.

[22] A. Goldberg and J. Hartline. Competitiveness via Consensus. In *Proceedings of the 14th ACM-SIAM Symposium on Discrete Algorithms*, pages 215 – 222, 2003.

[23] A. Goldberg and J. Hartline. Collusion-Resistant Mechanisms for Single-Parameter Agents. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms*, pages 620 – 629, 2005.

[24] A. Goldberg, J. Hartline, A. Karlin, and M. Saks. A Lower Bound on the Competitive Ratio of Truthful Auctions. In *Proceedings 21st Symposium on Theoretical Aspects of Computer Science*, pages 644–655, 2004.

[25] A. Goldberg, J. Hartline, A. Karlin, M. Saks, and A. Wright. Competitive Auctions and Digital Goods. *Games and Economic Behavior*, 2006.

[26] A. Goldberg, J. Hartline, and A. Wright. Competitive Auctions and Digital Goods. In *Proceeding of the 12th ACM-SIAM Symposium on Discrete Algorithms*, pages 735–744, 2001.

[27] A. Grigoriev, J. van Loon, R. Sitters, and M. Uetz. How to Sell a Graph: Guideliness for Graph Retailers. Meteor Research Memorandum RM/06/001, Maastricht University, 2005.

[28] V. Guruswami, J. Hartline, A. Karlin, D. Kempe, C. Kenyon, and F. McSherry. On Profit-Maximizing Envy-Free Pricing. In *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms*, pages 1164 – 1173, 2005.

[29] J. Hartline and V. Koltun. Near-Optimal Pricing in Near-Linear Time. In *Proceedings of the 9th Workshop on Algorithms and Data Structures*, pages 422–431, 2005.

[30] M. Kearns and U. Vazirani. *An Introduction to Computational Learning Theory.* MIT Press, 1994.

[31] V. Koltchinskii. Rademacher Penalties and Structural Risk Minimization. *IEEE Transactions of Information Theory*, 54(3):1902–1914, 2001.

[32] A. Likhodedov and T. Sandholm. Approximating Revenue-Maximizing Combinatorial Auctions. In *The Twentieth National Conference on Artificial Intelligence (AAAI)*, pages 267–274, 2005.

[33] R. Myerson. Optimal Auction Design. *Mathematics of Opperations Research*, 6:58–73, 1981.

[34] N. Nisan. Personal communication, 2005.

[35] N. Nisan. Introduction to mechanism design (for computer scientists). In N. Nisan, T. Roughgarden, E. Tardos, and V.V. Vazirani, editors, *Algorithmic Game Theory*, chapter 9. Cambridge University Press, 2007.

[36] V. Vapnik. *Statistical Learning Theory.* Springer-Verlag, 1998.

A Concentration Inequalities

Here is the McDiarmid inequality (see [16]) we use in our proofs:

Theorem 26 *Let $Y_1, ..., Y_n$ be independent random variables taking values in some set A, and assume that $t : A^n \rightarrow R$ satisfies:*

$$\sup_{y_1,...,y_n \in A, \overline{y}_i \in A} |t(y_1, ..., y_n) - t(y_1, ..., y_{i-1}, \overline{y}_i, y_{i+1}, y_n)| \leq c_i,$$

for all i, $1 \leq i \leq n$. Then for all $\gamma > 0$ we have:

$$\Pr\left\{|t(Y_1, ..., Y_n) - \mathbf{E}[t(Y_1, ..., Y_n)]| \geq \gamma\right\} \leq 2e^{-2\gamma^2 / \sum\limits_{i=1}^{n} c_i^2}$$

Here is also a consequence of the Chernoff bound that we used in Lemma 14.

Theorem 27 *Let $X_1, ..., X_n$ be independent Poisson trials such that, for $1 \leq i \leq n$, $\Pr[X_i = 1] = \frac{1}{2}$ and let $X = \sum_{i=1}^{n} X_i$. Then any n' we have:*

$$\Pr\left\{\left|X - \frac{n}{2}\right| \geq \epsilon \max\{n, n'\}\right\} \leq 2e^{-2n'\epsilon^2}$$