Revenue Maximization with a Single Sample

Peerapong Dhangwatnotai*
Department of Computer Science, Stanford University
460 Gates Building, 353 Serra Mall, Stanford, CA
pdh@cs.stanford.edu

Tim Roughgarden†
Department of Computer Science, Stanford University
462 Gates Building, 353 Serra Mall, Stanford, CA
tim@cs.stanford.edu

Qiqi Yan‡
Department of Computer Science, Stanford University
460 Gates Building, 353 Serra Mall, Stanford, CA
qiqiyan@cs.stanford.edu

ABSTRACT

We design and analyze approximately revenue-maximizing auctions in general single-parameter settings. Bidders have publicly observable attributes, and we assume that the valuations of indistinguishable bidders are independent draws from a common distribution. Crucially, we assume all valuation distributions are a priori *unknown* to the seller. Despite this handicap, we show how to obtain approximately optimal expected revenue — nearly as large as what could be obtained if the distributions were *known* in advance — under quite general conditions.

Our most general result concerns arbitrary downward-closed single-parameter environments and valuation distributions that satisfy a standard hazard rate condition. We also assume that no bidder has a unique attribute value, which is obviously necessary with unknown and attribute-dependent valuation distributions. Here, we give an auction that, for every such environment and unknown valuation distributions, has expected revenue at least a constant fraction of the expected optimal welfare (and hence revenue). A key idea in our auction is to associate each bidder with another that has the same attribute, with the second bidder's valuation acting as a random reserve price for the first. Conceptually, our analysis shows that even a single sample from a distribution — the second bidder's valuation — is sufficient information to obtain near-optimal expected revenue, even in quite general settings.

Categories and Subject Descriptors

F.0 [**Theory of Computation**]: General

*Supported in part by the ONR Young Investigator Award of the second author.

†Supported in part by NSF CAREER Award CCF-0448664, an AFOSR MURI grant, an ONR Young Investigator Award, an ONR PECASE Award, and an Alfred P. Sloan Fellowship.

‡Supported by a Stanford Graduate Fellowship.

EC'10, June 7–11, 2010, Cambridge, Massachusetts, USA.

General Terms

Economics, Theory, Algorithms

Keywords

Mechanism Design, Auctions, Revenue Maximization

1. INTRODUCTION

Worst-Case Revenue Maximization.

Revenue guarantees for auctions are an important but elusive goal. A key challenge, first identified by Goldberg et al. [6], is to develop a useful competitive analysis framework. The obvious approach, familiar from e.g. online algorithms, is to define the "optimal" revenue as that achievable by an all-knowing adversary — an entity who, unlike the auctioneer, knows all of the bidders' private information (like valuations) a priori. Such an opponent turns out to be simply too powerful: even in very simple problems, no incentive-compatible auction can obtain more than a vanishingly small fraction of its revenue in the worst case.

A successful competitive analysis framework must, obviously, differentiate between intuitively "better" auctions and "worse" ones. Goldberg et al. [6] proposed a *revenue benchmark* approach, which has been applied successfully to a number of auction settings (see [8] for a survey). The idea is to define a real-valued function on inputs (i.e., bid vectors) that represents an upper bound on the maximum revenue achievable by any "reasonable" auction on each input. Such benchmarks are generally smaller than the revenue achievable by an all-knowing seller, opening up the possibility of non-trivial (but still well motivated) worst-case revenue guarantees. The best known such benchmark is $\mathcal{F}_2$ for digital goods (i.e., unlimited supply and unit-demand) auctions, which for a given set of bids is defined as the maximum revenue achievable using a common selling price and selling to at least two bidders [6].

Researchers have successively designed novel auctions that are competitive with benchmarks like $\mathcal{F}_2$ — i.e., auctions that obtain at least a constant fraction of the benchmark revenue on every possible input — for several auction settings (see [8]). Thus far, however, most constant-factor approximations in the revenue benchmark framework have been confined to simple auction settings, where the goods are in unlimited supply and/or the bidders are symmetric. For example, no worst-case revenue guarantees of any sort are known for the central problem of combinatorial auctions with (known) single-minded bidders.[1]

[1]In such an auction, there are n bidders and m goods with

In the classical auction literature, these issues are skirted by assuming that the seller knows, a priori, a distribution over the private information of the bidders. Assuming such a distribution over inputs obviates the need for a competitive analysis framework: since the performance of every auction can be summarized by a single number — its expected revenue with respect to the assumed distribution over inputs — there is an auction that is unequivocally optimal. For example, Myerson [16] characterizes the optimal auctions for every single-parameter setting (including, e.g., combinatorial auctions with single-minded bidders) and every product distribution over the bidders' parameters. The obvious criticisms of this approach are that: (1) private information is assumed to come from a distribution; (2) the precise distribution is assumed to be known to the seller. Of particular concern are auction design results that depend on the details of the assumed distribution, as then modest uncertainty about the distribution translates to significant uncertainty about what auction to use. The natural goal of "detail-free" auctions, whose structure is largely independent of the details of the distribution, is well known in economics and is often called "Wilson's Doctrine" [20].

Our approach.

We propose an analysis framework that combines the advantages of both the revenue benchmark approach (detail-free auctions and robust approximation guarantees) and the classical economics approach (results for general single-parameter problems). The framework is natural: we consider environments in which private information is drawn from a distribution (retaining assumption (1), above) but in which this distribution is a priori *unknown* to the seller (discarding assumption (2)).

> *The goal is to design a single auction such that, whatever the underlying distribution, its expected revenue is almost as large as that of an optimal auction tailored for that distribution.*

Our Results.

Our primary contribution is the *Single Sample* mechanism, and a proof that it *simultaneously approximates* all Bayesian-optimal mechanisms for all valuation distributions under reasonably general assumptions. In more detail, we consider binary single-parameter environments, where the feasible outcomes are described by a collection of bidder subsets. For example, in combinatorial auctions with single-minded bidders, feasible subsets correspond to bidders seeking mutually disjoint bundles of goods. Each bidder has a private valuation for belonging to the chosen feasible set. Bidders can be asymmetric, in that each bidder has an observable attribute, and we assume that the valuations of bidders with a common attribute are drawn i.i.d. from an (unknown) distribution that satisfies standard technical conditions (see Section 2). Bidders with different attributes can have valuations drawn (independently) from completely different distributions. For example, based on (publicly observable) eBay bidding history, one might classify bidders into "bargain-hunters", "typical", and "aggressive", with the expectation that bidders in the same class are likely to bid similarly (without necessarily knowing what their valuations for a given item might be). Simple and standard examples imply that non-trivial simultaneous approximation guarantees with unknown distributions are possible only when the environment is *non-singular*, meaning that there is no bidder with a unique attribute (see e.g. [6]).

Precisely, our most general result is a mechanism that, for every non-singular downward-closed environment and attribute-dependent valuation distributions that satisfy a standard hazard rate condition, has expected revenue at least a constant fraction of the expected optimal welfare (and hence revenue) in that environment. This gives, as an example special case, the first revenue guarantee for combinatorial auctions with single-minded bidders outside of the standard Bayesian setup with known distributions [11, 13]. We prove an approximation guarantee of $\frac{1}{4} \cdot \frac{\kappa-1}{\kappa}$ when there are at least $\kappa \geq 2$ bidders with each attribute, and our analysis of our mechanism is tight (for a worst-case distribution) for each κ. This approximation bound depends on solving the underlying welfare-maximization problem exactly — which is NP-hard in some contexts — but our design and analysis techniques easily extend to approximate welfare-maximization algorithms.

Second, we prove better approximation bounds for *i.i.d. matroid* environments, where all bidders have the same attribute and the feasible subsets form a matroid on the set of bidders. Examples of such environments include k-unit auctions and certain matching markets. Here, we prove an approximation factor of $\frac{1}{2}$ relative to the expected revenue of an optimal mechanism, for every $\kappa \geq 2$. Moreover, in this result we only require that the valuation distribution is "regular", a condition that is weaker than the hazard rate condition above and permits distributions with heavier tails.

Third, we provide better approximation guarantees when κ is large. Specifically, for every $\epsilon > 0$, when $\kappa = \Omega(\text{poly}(\epsilon^{-1}))$, we show how to achieve a $(1-\epsilon)$-approximation of the optimal expected revenue in every i.i.d. matroid environment and a $(1-\epsilon)\frac{1}{e}$-approximation of the optimal expected welfare in every downward-closed environment, under the same distributional assumptions as above. (Here e denotes $2.718\ldots$.) Our lower bound on κ depends *only* on the approximation parameter ϵ and *not* on the underlying valuation distributions.

The Main Ideas.

A key idea in our Single Sample mechanism is to associate each bidder with another that has the same attribute, with the second bidder's valuation acting as a random reserve price for the first. Conceptually, our analysis shows that even a single sample from a distribution — the second bidder's valuation — is sufficient information to obtain near-optimal expected revenue, even in quite general settings.

For a single-item auction, the Bulow-Klemperer theorem[2] furnishes good intuition for why a random reserve price might be an effective surrogate for an optimal one. It is not at all obvious, however, that this intuition should extend to more complex problems in which the bidders are not interchangeable — that a "local" guarantee for a single bidder

only one unit of each. Each bidder i wants a particular subset S_i of goods (e.g., a set of geographically clustered wireless spectrum licenses) and has a valuation v_i for it. The seller knows the set S_i a priori but not the valuation.

[2]The Bulow-Klemperer theorem [3] states: for every $n \geq 1$ and valuation distribution F that is regular in the sense of Section 2, the expected revenue of the Vickrey auction with $n+1$ bidders with valuations drawn i.i.d. from F is at least that of a revenue-maximizing auction with n such bidders.

should automatically extend to a "global" guarantee about the expected revenue of an entire allocation computed using a collection of random reserves. In fact, this "local to global" translation does *not* hold in arbitrary downward-closed environments with regular valuations distributions: our Single Sample mechanism does not always give a constant-factor approximation in such settings. Our proofs implement this "local to global" approach by making careful use of additional problem structure, either in the feasible sets or in the valuation distributions.

Our final result gives an asymptotically optimal revenue guarantee as the number of bidders of every attribute tends to infinity, A weak version of this result, which does not give quantitative bounds on the number of bidders required, can be derived from the Law of Large Numbers. To prove our distribution-independent polynomial bound on the number of bidders needed, we show that there exists a set of "quantiles" that is simultaneously small enough that concentration bounds can be usefully applied, and rich enough to guarantee a good approximation for every regular valuation distribution. Our approximation bound relies on a geometric characterization of regular distributions.

Further Related Work.

Previous works that can be interpreted as simultaneous approximation for multiple distributions include Segal [19] and Neeman [17] (who consider asymptotic optimality for symmetric bidders and identical goods), and the result of Bulow and Klemperer [3] mentioned above. Hartline and Roughgarden [11] extended the Bulow-Klemperer result (with a small approximation loss and slightly stronger technical conditions) to general single-parameter settings.

Approximation in the revenue benchmark framework discussed in the introduction is strictly stronger than the simultaneous approximation goal pursued in the present paper; this fact is made explicit in [10] and is pursued further in [4, 11, 12]. The point of this paper is to obtain simultaneous approximation w.r.t. a class of distributions for problems much more general than those studied in [4, 10, 11, 12] — which are confined to problems with symmetric bidders and/or environments — and via simpler mechanisms, and with better approximation guarantees. For example, our simple Single Sample mechanism is reminiscent of the "Pairing Mechanism" studied in [7], but the latter is not constant-competitive in the revenue benchmark analysis framework. For a different example, for simultaneous approximation in digital goods auctions with an unknown i.i.d. regular valuation distribution, we achieve an approximation factor of $\frac{1}{2}$ (Remark 3.7); in the revenue benchmark analog of this problem (a worst-case approximation of the benchmark $\mathcal{F}_2$), no truthful auction has an approximation factor better than .42 [6].

2. PRELIMINARIES

This section reviews standard terminology and facts about Bayesian-optimal mechanism design. We encourage the reader familiar with these to skip to Section 3.

Environments.

An environment is defined by a set E of bidders, and a collection $\mathcal{I} \subseteq 2^E$ of feasible sets of bidders, which are the subsets of bidders that can simultaneously "win". We always assume that the set system $(E, \mathcal{I})$ is *downward-closed*, meaning that if $T \in \mathcal{I}$ and $S \subseteq T$ then $S \in \mathcal{I}$. Each bidder has a publicly observable *attribute* drawn from a known set A. We assume that each bidder with attribute a has a private *valuation* for winning that is an independent draw from a distribution F_a. We sometimes denote an environment by a tuple $Env = (E, \mathcal{I}, A, (a_i)_{i \in E}, (F_a)_{a \in A})$. Every subset $T \subseteq E$ of bidders induces a subenvironment in a natural way, with feasible sets $\{S \cap T\}_{S \in \mathcal{I}}$.

We classify an environment according to its set system, the attributes of its bidders, and its underlying valuation distributions. In a *matroid environment*, the system $(E, \mathcal{I})$ is a matroid.[3] Examples include digital goods (where $\mathcal{I} = 2^E$), k-unit auctions (where $\mathcal{I}$ is all subsets of size at most k), and certain unit-demand matching markets (corresponding to a transversal matroid). An environment is *non-singular* if there is no bidder with a unique attribute, and is *i.i.d.* if every bidder has the same attribute. An environment is *regular* or *m.h.r.* if every valuation distribution is a regular distribution or an m.h.r. distribution (as defined below), respectively.

Truthful Mechanisms.

Name the bidders $E = \{1, 2, \ldots, n\}$. A (deterministic) mechanism $\mathcal{M}$ comprises an *allocation rule* $\mathbf{x}$ that maps every bid vector $\mathbf{b}$ to a characteristic vector of a feasible set (in $\{0,1\}^n$), and a *payment rule* $\mathbf{p}$ that maps every bid vector $\mathbf{b}$ to a non-negative payment vector in $[0, \infty)^n$. We insist on individual rationality in the sense that $p_i(\mathbf{b}) \leq b_i \cdot x_i(\mathbf{b})$ for every i and $\mathbf{b}$. We assume that each bidder i aims to maximize the quasi-linear utility $u_i(\mathbf{b}) = v_i \cdot x_i(\mathbf{b}) - p_i(\mathbf{b})$, where v_i is its private valuation for winning. We say $\mathcal{M}$ is *truthful* if for every bidder i and fixed bids $\mathbf{b}_{-i}$ of the others, bidder i maximizes its utility by setting its bid b_i to its private valuation v_i. Since we only consider truthful mechanisms, in the rest of the paper we use valuations and bids interchangeably. A well-known characterization of truthful mechanisms in single-parameter settings [16, 1] says that a mechanism $(\mathbf{x}, \mathbf{p})$ is truthful if and only if the allocation rule is monotone — $x_i(b'_i, \mathbf{b}_{-i}) \geq x_i(\mathbf{b})$ for every i, $\mathbf{b}$, and $b'_i \geq b_i$ — and the payment rule is given by a certain formula of the allocation rule. We often specify a truthful mechanism by its monotone allocation rule, with the understanding that it is supplemented with the unique payment rule that yields a truthful mechanism.

For example, the *VCG mechanism*, which chooses the feasible set S that maximizes the welfare $\sum_{i \in S} v_i$, has a monotone allocation rule and can be made truthful using suitable payments. Two variants of the VCG mechanism are also important in this paper. Let r_i be a *reserve price* for bidder i. The *VCG mechanism with eager reserves* $\mathbf{r}$ works as follows, given bids $\mathbf{v}$: (1) delete all bidders i with $v_i < r_i$; (2) run the VCG mechanism on the remaining bidders to determine the winners; (3) charge each winning bidder i the larger of r_i and its VCG payment in step (2). In the *VCG mechanism with lazy reserves* $\mathbf{r}$, steps (1) and (2) are reversed. Both of these mechanisms are feasible and truthful in every downward-closed environment. The two variants are equivalent in sufficiently simple environments (as we show in Corollary 3.4), but are different in general.

The *efficiency* of *welfare* of the outcome of a mechanism

[3]Recall that a *matroid* (e.g. [18]) is a ground set E and a non-empty downward-closed collection $\mathcal{I} \subseteq 2^E$ of *independent sets* such that whenever $S, T \in \mathcal{I}$ with $|T| < |S|$, there is some $i \in S \setminus T$ such that $T \cup \{i\} \in \mathcal{I}$.

is the sum of the winners' valuations, and the *revenue* is the sum of the winners' payments (which can only be less). We write $Eff_{\mathcal{M}}(Env)$ and $Rev_{\mathcal{M}}(Env)$ for the expected efficiency and expected revenue of the mechanism $\mathcal{M}$ in the environment Env, respectively, where the expectation is over the random bidder valuations.

Bayesian-Optimal Auctions.

Let F be (the cumulative distribution function of) a distribution. For simplicity, we assume that the distribution has a finite support in the form of a closed interval $[l, h]$, and has a positive and smooth density function. When convenient, we assume that $l = 0$; a simple "shifting argument" shows that this is the worst case for approximate revenue guarantees. The *virtual valuation function* w.r.t. F is defined as $\varphi_F(v) = v - 1/h(v)$, where $h(v) = \frac{f(v)}{1-F(v)}$ is the *hazard rate function* of F. This paper works with two different assumptions on valuation distributions. A *regular* distribution has, by definition, a nondecreasing virtual valuation function. A *monotone hazard rate (m.h.r.)* distribution has a nondecreasing hazard rate function. Many important distributions (exponential, uniform, Gaussian, etc.) are m.h.r.; intuitively, these are distributions with tails no heavier than the exponential distribution. Regular distributions include all m.h.r. distributions along with some additional distributions with heavier tails (e.g., some power-law distributions).

Myerson [16] characterized the expected revenue-maximizing mechanisms for single-parameter environments using the following key lemma.

Lemma 2.1 (Myerson's Lemma) *For every truthful mechanism* $(\mathbf{x}, \mathbf{p})$*, the expected payment of a bidder* i *with valuation distribution* F_i *satisfies*

$$E_{\mathbf{v}}[p_i(\mathbf{v})] = E_{\mathbf{v}}[\varphi_{F_i}(v_i) \cdot x_i(\mathbf{v})].$$

Moreover, this identity holds even after conditioning on the bids $\mathbf{v}_{-i}$ *of the bidders other than* i.

In words, the (conditional) expected payment of a bidder is precisely its (conditional) expected contribution to the virtual welfare. It follows that if the distributions are regular, then a revenue-maximizing truthful mechanism chooses a feasible set S that maximizes the virtual welfare $\sum_{i \in S} \varphi_{F_i}(v_i)$. (The role of regularity is to ensure that this allocation rule is indeed monotone; otherwise, additional ideas are needed [16].)

3. REVENUE GUARANTEES WITH A SINGLE SAMPLE

In this section we design a prior-free auction that simultaneously approximates the optimal expected revenue to within a constant factor in every non-singular m.h.r. single-parameter environment. Section 3.1 defines our mechanism. Section 3.2 introduces some of our main analysis techniques in the simpler setting of matroid environments with i.i.d. valuation distributions, where we can also obtain better approximation bounds than in the general case. Section 3.3 proves our main result for general downward-closed environments. Section 3.5 notes that our analysis carries over easily to computationally efficient variants of our mechanism.

3.1 The Single Sample Mechanism

We analyze the following mechanism.

Definition 3.1 (Single Sample) Given a non-singular downward-closed environment $Env = (E, \mathcal{I}, A, (a_i)_{i \in E}, (F_a)_{a \in A})$, the *Single Sample* mechanism is the following:

(1) For each represented attribute a, pick a *reserve bidder* i_a with attribute a uniformly at random from all such bidders.

(2) Run the VCG mechanism on the sub-environment induced by the non-reserve bidders to obtain a preliminary winning set P.

(3) For each bidder $i \in P$ with attribute a, place i in the final winning set W if and only if $v_i \geq v_{i_a}$. Charge every winner $i \in W$ with attribute a the maximum of its VCG payment computed in step (2) and the reserve price v_{i_a}.

In other words, we randomly pick one bidder of each attribute to set a reserve price for the other bidders with that attribute, and then run the VCG mechanism with lazy reserves on the remaining bidders. The Single Sample mechanism is clearly prior-free — that is, it is defined independently of the F_a's — and it is easy to verify that it is truthful.

3.2 Warm-Up: I.I.D. Matroid Environments

To illustrate some of our main techniques in a relatively simple setting, we first consider matroid environments in which all bidders have the same attribute (i.e., have i.i.d. valuations). For such settings, we only need to assume that the common valuation distribution F is regular (recall Section 2 for definitions).

Theorem 3.2 (I.I.D. Matroid Environments) *For every i.i.d. regular matroid environment with at least* $n \geq 2$ *bidders, the expected revenue of the Single Sample mechanism is at least a* $\frac{1}{2} \cdot \frac{n-1}{n}$ *fraction of that of an optimal mechanism for the environment.*

The factor of $(n-1)/n$ can be removed with a minor tweak to the mechanism (Remark 3.7). Section 4 considers the case of large n and shows how to use multiple samples to obtain better approximation factors.

What's so special about i.i.d. regular matroid environments? Define a *monopoly reserve price* of a valuation distribution F as a price in $\text{argmax}_p[p \cdot (1 - F(p))]$. Then, Myerson's Lemma easily implies the following.

Proposition 3.3 (E.g. [5]) *In every i.i.d. regular matroid environment, the VCG mechanism with eager monopoly reserves is a revenue-maximizing mechanism.*

The matroid assumption also allows us to pass from eager to lazy reserves.

Corollary 3.4 *In every i.i.d. regular matroid environment, the VCG mechanism with lazy monopoly reserves is a revenue-maximizing mechanism.*

PROOF. The VCG mechanism can be implemented in a matroid environment via the greedy algorithm: bidders are considered in nonincreasing order of valuations, and a bidder is added to the winner set if and only if doing so preserves feasibility (given the previous selections). With a common reserve price r, it makes no difference whether bidders with valuations below r are thrown out before or after running the greedy algorithm. Thus in matroid environments, the VCG mechanism with a lazy common reserve is equivalent to the VCG mechanism with an eager common reserve. □

Proving an approximate revenue-maximization guarantee for the Single Sample mechanism thus boils down to understanding the two ways in which it differs from the VCG mechanism with lazy monopoly reserves — it throws away a random bidder, and it uses a random reserve rather than a monopoly reserve. The damage from the first difference is easy to control.

Lemma 3.5 *In expectation over the choice of the reserve bidder, the expected revenue of an optimal mechanism for the environment induced by the non-reserve bidders is at least an $(n-1)/n$ fraction of the expected revenue of an optimal mechanism for the original environment.*

Proof. Condition first on the valuations of all bidders and let S denote the winners under the optimal mechanism for the full environment. Since the reserve bidder is chosen independently of the valuations, each bidder of S is a non-reserve bidder with probability $(n-1)/n$. By the linearity of expectation, the expected virtual welfare — over the choice of the reserve bidder, with all valuations fixed — of the (feasible) set of non-reserve bidders of S is $(n-1)/n$ times that of S, and the expected maximum-possible virtual welfare in the sub-environment is at least this. Taking expectations over bidder valuations and applying Myerson's Lemma (Lemma 2.1) completes the proof. □

The crux of the proof of Theorem 3.2 is to show that a random reserve price serves as a sufficiently good approximation of a monopoly reserve price. The next key lemma formalizes this goal for the case of a single bidder. Its proof uses a geometric property of regular distributions. To explain it, for a regular distribution F, define the *normalized revenue function* R as $R(q) = q \cdot F^{-1}(1-q)$ for all $q \in [0,1]$. Here, q represents the probability of a sale to a bidder with a valuation drawn from F, and $R(q)$ the corresponding expected revenue. We also define the *revenue function* by $\hat{R}(p) = p(1-F(p))$, which is the same quantity parameterized by the selling price p. That is, $\hat{R}(p)$ is the expected revenue of a single-item auction with posted price p and a single bidder with valuation drawn from F. An example of a normalized revenue function is shown in Figure 1. One can check easily that the derivative $R'(q)$ equals the virtual valuation $\varphi_F(p)$, where $p = F^{-1}(1-q)$. Regularity of F thus implies that $R'(q)$ is nonincreasing and hence R is concave. Also, assuming that the support of F is $[0,h]$ for some $h > 0$ — recall Section 2 — we have $R(0) = R(1) = 0$.

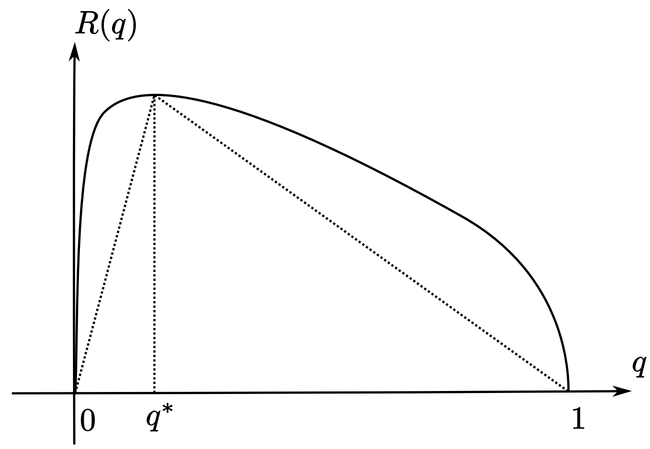

Figure 1: The normalized revenue function of a regular distribution.

Lemma 3.6 *Let F be a regular distribution with monopoly price r^* and revenue function $\hat{R}$. Let v denote a random valuation from F. For every nonnegative number $t \geq 0$,*

$$\mathbf{E}_v\left[\hat{R}(\max\{t,v\})\right] \geq \frac{1}{2} \cdot \hat{R}(\max\{t,r^*\}). \quad (1)$$

Proof. For intuition, first suppose that $t = 0$. Then the claim is equivalent to the assertion that the expectation of $R(q)$ is at least half of $R(q^*)$, where q and q^* solve $q = 1-F(v)$ and $q^* = 1-F(r^*)$, respectively. Since q is uniformly distributed on $[0,1]$, $\mathbf{E}_q[R(q)]$ equals the area under the curve defined by R. By concavity, this area is at least the area of the triangle in Figure 1, which is $\frac{1}{2} \cdot 1 \cdot R(q^*) = \frac{1}{2} \cdot \hat{R}(r^*)$.

If $0 < t < r^*$, then the right-hand side of (1) is unchanged while the left-hand side of (1) only increases — the only difference is to sometimes use a selling price t that is better than the previous selling price v. (We are using concavity of the revenue curve here.) Finally, if $t > r^*$, then the right-hand side of (1) is $\hat{R}(t)$; and the left-hand side is a convex combination of $\hat{R}(t)$ (when $v \leq t$) and the expected value of $\hat{R}(q)$ when q is drawn uniformly from $[t,1]$, which by concavity is at least $\hat{R}(t)/2$. □

We now prove Theorem 3.2 by extending the approximation bound in Lemma 3.6 from a single bidder to all bidders and blending in Lemma 3.5.

Proof of Theorem 3.2: Condition on the choice of the reserve bidder j. Fix a non-reserve bidder i and condition on all valuations except those of i and j. This is enough information to uniquely determine the VCG threshold $t(\mathbf{v}_{-i})$ for i. (Recall that j, as a reserve bidder, does not participate in the VCG computation in step (2) of the Single Sample mechanism.) After this conditioning, we can analyze bidder i as in a single-bidder auction, with an extra external reserve price of $t(\mathbf{v}_{-i})$. Let r^* and $\hat{R}$ denote the monopoly price and revenue function for the underlying regular distribution F, respectively. The conditional expected revenue that i contributes to the revenue-maximizing solution in the sub-environment of non-reserve bidders is $\hat{R}(\max\{t(\mathbf{v}_{-i}), r^*\})$. The conditional expected revenue that i contributes to the Single Sample mechanism is $\mathbf{E}_{v_j}\left[\hat{R}(\max\{t(\mathbf{v}_{-i}), v_j\})\right]$. Since v_i, v_j are independent samples from the regular distribution F, Lemma 3.6 implies that the latter conditional expectation is at least 50% of the former. Taking expectations over the previously fixed valuations of bidders other than i and j, summing over the non-reserve bidders i and applying linearity of expectation, and finally taking the expectation over the choice of the reserve bidder j and applying Lemma 3.5 proves the theorem. □

Remark 3.7 (Optimized Version of Theorem 3.2) We can improve the approximation guarantee in Theorem 3.2 from $\frac{1}{2} \cdot \frac{n-1}{n}$ to $\frac{1}{2}$. Instead of discarding the reserve bidder j, we include it in the VCG computation in step (2) of the Single Sample mechanism. An arbitrary other bidder h is used to set a reserve price v_h for the reserve bidder j. Like the other bidders, the reserve bidder is included in the final winning set W if and only if it is chosen by the VCG mechanism in step (2) and also has a valuation above its reserve price ($v_j \geq v_h$). Its payment is then the maximum of its VCG payment and v_h.

The key observation is that, for every choice of a reserve bidder j, a non-reserve bidder i, and valuations $\mathbf{v}$, bidder i wins with bidder j included in the VCG computation in step (2) if and only if it wins with bidder j excluded from the computation. Like Corollary 3.4, this observation can be derived from the fact that the VCG mechanism can be implemented via a greedy algorithm in i.i.d. regular matroid environments. If $v_i \leq v_j$, then i cannot win in either case (it fails to clear the reserve); and if $v_i > v_j$, then the greedy algorithm considers bidder i before j even if the latter is included in the VCG computation.

Thus, the expected revenue from non-reserve bidders is the same in both versions of the Single Sample mechanism. In the modified version, the obvious analog of Lemma 3.5 for a single bidder and Lemma 3.6 imply that the reserve bidder also contributes, in expectation, a $\frac{1}{2} \cdot \frac{1}{n}$ fraction of the expected revenue of an optimal mechanism. Combining the contributions of the reserve and non-reserve bidders yields an approximation guarantee of $\frac{1}{2}$ for the modified mechanism. Simple examples show that this analysis, and hence also the bound in Lemma 3.6, is tight in the worst case.

3.3 General Downward-Closed Environments

We now give our main result for the performance of the Single Sample mechanism in general downward-closed environments. We first note that, unlike in the special case of matroid environments, the mechanism does not admit a constant-factor approximate revenue guarantee for every i.i.d. regular environment; see the full version for details.

Statement of Main Result.

Duly warned by the example mentioned above, we consider only (not necessarily identical) m.h.r. valuation distributions in this section. Our main result is a constant-factor guarantee for the Single Sample mechanism in arbitrary nonsingular m.h.r. environments, even with respect to the expected optimal *welfare*.

Theorem 3.8 (General Environments) *For every m.h.r. environment with at least $\kappa \geq 2$ bidders of every present attribute, the expected revenue of the Single Sample mechanism is at least a $\frac{1}{4} \cdot \frac{\kappa-1}{\kappa}$ fraction of the expected optimal welfare in the environment.*

The bound in Theorem 3.8 is $1/8$ when κ is its minimum allowable value of 2, and it converges quickly to $1/4$ as κ grows. Our analysis of the Single Sample mechanism is tight for all values of $\kappa \geq 2$, as shown by a digital goods environment with κ bidders with valuations drawn i.i.d. from an exponential distribution with rate 1: the expected optimal welfare is κ, and a calculation shows that the expected revenue of Single Sample is $(\kappa - 1)/4$.

Since welfare obviously upper bounds the revenue obtainable by any mechanism, we have the following corollary.

Corollary 3.9 *For every m.h.r. environment with at least $\kappa \geq 2$ bidders of every present attribute, the expected revenue of the Single Sample mechanism is at least a $\frac{1}{4} \cdot \frac{\kappa-1}{\kappa}$ fraction of that of the optimal mechanism for the environment.*

A Weaker Result via VCG with Lazy Monopoly Reserve Prices.

General downward-closed environments pose a number of challenges absent from i.i.d. matroid environments. The expected revenue-maximizing mechanism is generally complicated — nothing as simple as the VCG mechanism with suitable reserve prices. Eager and lazy reserve prices are no longer equivalent, and the lazy reserve prices in the Single Sample mechanism are crucial in our analysis.

The simplest approach to establishing a constant-factor approximation guarantee for the Single Sample mechanism is to prove that the VCG mechanism with lazy monopoly reserves is a reasonable approximation of an optimal mechanism, and then proceed as in the previous section. We prove such a result next. We believe that this is interesting in its own right, and it gives a slightly weaker version of Theorem 3.8 as a corollary. We then prove the better bound claimed in the theorem via a more delicate argument.

One ingredient of the analysis of the previous section carries over easily to the present one — controlling the expected lost welfare from discarded reserve bidders.

Lemma 3.10 *The expected optimal welfare in the sub-environment induced by non-reserve bidders is at least a $(\kappa - 1)/\kappa$ fraction of that in the original environment.*

The proof of Lemma 3.10 is essentially the same as that of Lemma 3.5, with valuations assuming the role previously played by virtual valuations. In contrast to Remark 3.7, discarding reserve bidders before the VCG computation in step (2) is important for the analysis of the Single Sample mechanism in non-matroid environments.

As in the previous section, we require a technical lemma about the single-bidder case.

Lemma 3.11 *Let F be an m.h.r. distribution with monopoly price r^* and revenue function $\hat{R}$. Let $V(t)$ denote the expected welfare of a single-item auction with a posted price of t and a single bidder with valuation drawn from F. For every nonnegative number $t \geq 0$,*

$$\hat{R}(\max\{t, r^*\}) \geq \frac{1}{e} \cdot V(t). \quad (2)$$

Proof. Let s denote $\max\{t, r^*\}$. Recall that, by the definition of the hazard rate function, $1 - F(x) = e^{-H(x)}$ for every $x \geq 0$, where $H(x)$ denotes $\int_0^x h(z)dz$. Note that since $h(z)$ is non-negative and nondecreasing, $H(x)$ is nondecreasing and convex. We can write the left-hand side of (2) as $s \cdot (1 - F(s)) = s \cdot e^{-H(s)}$ and, for a random sample v from F,

$$\begin{aligned} V(t) &= \mathbf{Pr}[v \geq t] \cdot \mathbf{E}[v \mid v \geq t] \\ &= e^{-H(t)} \cdot \left[t + \int_t^\infty e^{-(H(v) - H(t))} dv \right]. \quad (3) \end{aligned}$$

By convexity of the function H, we can lower bound its value using a first-order approximation at s:

$$H(v) \geq H(s) + H'(s)(v-s) = H(s) + h(s)(v-s) \quad (4)$$

for every $v \geq 0$. There are now two cases. If $t \leq r^* = s$, then $h(s) = 1/s$ since r^* is a monopoly price.[4] Starting from (3) and using that H is nondecreasing, and then substituting (4) yields

$$\begin{aligned} V(t) &\leq \int_0^\infty e^{-H(v)} dv \\ &\leq \int_0^\infty e^{-(H(s)+\frac{v-s}{s})} dv \\ &= e \cdot s \cdot e^{-H(s)}. \end{aligned}$$

If $r^* \leq t = s$, then the m.h.r. assumption implies that $h(s) \geq 1/s$ and (4) implies that $H(v) \geq H(t) + (v-t)/t$ for all $v \geq t$. Substituting into (3) gives

$$\begin{aligned} V(t) &\leq e^{-H(t)} \cdot \left[t + \int_t^\infty e^{-(H(t)+\frac{v-t}{t}-H(t))} dv \right] \\ &\leq e^{-H(t)} \cdot \int_0^\infty e^{-\frac{v-t}{t}} dv \\ &= e \cdot s \cdot e^{-H(s)}, \end{aligned}$$

where in the second inequality we use that $e^{-(v-t)/t} \geq 1$ for every $v \leq t$. □

Lemma 3.11 implies that the expected revenue of the VCG mechanism with lazy reserve prices is competitive with the expected optimal welfare in every downward-closed environment with (not necessarily identical) m.h.r. valuation distributions.

Theorem 3.12 (VCG with Lazy Monopoly Reserves) *For every m.h.r. environment, the expected revenue of the VCG mechanism with lazy monopoly reserves is at least a $1/e$ fraction of the expected efficiency of the VCG mechanism.*

PROOF. Fix a bidder i and valuations $\mathbf{v}_{-i}$. This determines a winning threshold t for bidder i under the VCG mechanism (with no reserves). Lemma 3.11 implies that the conditional expected revenue obtained from i in the VCG mechanism with lazy monopoly reserves is at least a $1/e$ fraction of the conditional expected welfare obtained from i in the VCG mechanism (with no reserves). Taking expectations over $\mathbf{v}_{-i}$ and summing over all the bidders proves the theorem. □

Considering a single bidder with an exponentially distributed valuation shows that the bounds in Lemma 3.11 and Theorem 3.12 are tight in the worst case.

The arguments in Section 3.2 directly imply that the Single Sample mechanism obtains essentially half of the expected revenue of the VCG mechanism with lazy monopoly reserves. Mimicking the proof of Theorem 3.2, with Lemma 3.10 replacing Lemma 3.5, gives the following weaker version of Theorem 3.8.

[4]One proof of this follows from the first-order condition for the revenue function $p(1-F(p))$; alternatively, applying Myerson's Lemma to the single-bidder case shows that $r^* = \varphi_F^{-1}(0)$ and hence $r^* - 1/h(r^*) = \varphi_F(r^*) = 0$.

Theorem 3.13 (A Weaker Single Sample Guarantee) *For every m.h.r. environment with at least $\kappa \geq 2$ bidders of every present attribute, the expected revenue of the Single Sample mechanism is at least a $\frac{1}{2e} \cdot \frac{\kappa-1}{\kappa}$ fraction of the expected optimal welfare in the environment.*

Proof of Main Result (Theorem 3.8).

To obtain the better bound claimed in Theorem 3.8, we need to optimize jointly the two single-bidder guarantees in Lemmas 3.6 and 3.11. This is done in the next lemma.

Lemma 3.14 *Let F be an m.h.r. distribution with monopoly price r^* and revenue function $\hat{R}$, and define $V(t)$ as in Lemma 3.11. For every nonnegative number $t \geq 0$,*

$$\mathbf{E}_v\left[\hat{R}(\max\{t,v\})\right] \geq \frac{1}{4} \cdot V(t). \quad (5)$$

PROOF. Recall from the proof of Lemma 3.11 that $V(t)$ can be written as in (3); we show that the left-hand side of (5) is at least 25% of that quantity.

Consider two i.i.d. samples v_1, v_2 from F. We interpret v_2 as the random reserve price v in (5) and v_1 as the valuation of the single bidder. The left-hand side of (5) is equivalent to the expectation of a random variable that is equal to t if $v_2 \leq t \leq v_1$, which occurs with probability $F(t) \cdot (1-F(t))$; equal to v_2 if $t \leq v_2 \leq v_1$, which occurs with probability $\frac{1}{2}(1-F(t))^2$; and equal to zero, otherwise. Hence,

$$\begin{aligned} &\mathbf{E}_v\left[\hat{R}(\max\{t,v\})\right] \\ &\geq \frac{1}{2}\big(F(t)\cdot(1-F(t))\cdot t \quad (6) \\ &\quad +(1-F(t))^2 \cdot \mathbf{E}[\min\{v_1,v_2\} \mid \min\{v_1,v_2\} \geq t]\big) \\ &= \frac{1}{2}(1-F(t)) \cdot \Big(t \cdot F(t) + (1-F(t)) \cdot \Big[t \\ &\quad + e^{2H(t)} \int_t^\infty e^{-2H(v)} dv\Big]\Big) \\ &\geq \frac{1}{2}(1-F(t)) \cdot \left[t + e^{H(t)} \int_t^\infty e^{-H(2v)} dv\right] \quad (7) \\ &= \frac{1}{4}(1-F(t)) \cdot \left[2t + \int_{2t}^\infty e^{-(H(v)-H(t))} dv\right] \\ &\geq \frac{1}{4}(1-F(t)) \cdot \left[t + \int_t^\infty e^{-(H(v)-H(t))} dv\right], \quad (8) \end{aligned}$$

where in (7) and (8) we are using that H is non-negative, nondecreasing, and convex. Comparing (3) and (8) proves the lemma. □

The proof of Theorem 3.8 is the same as that of Theorem 3.2, with the following substitutions: the welfare of the VCG mechanism (with no reserves) plays the previous role of the revenue of the VCG mechanism with lazy monopoly reserves; Lemma 3.14 replaces Lemma 3.6; and Lemma 3.10 takes the place of Lemma 3.5.

3.4 Non-I.I.D. Matroid Environments

We can also use a different mechanism to obtain a constant fraction of the optimal expected revenue in matroid environments with valuation distributions that are not necessarily identical. We provide here only one simple approach, and defer proofs and optimized constants to the full version.

Definition 3.15 (Single Sample 2) Given a non-singular downward-closed environment Env, the *Single Sample 2 (SS2)* mechanism creates a new environment Env' by, for every attribute a, arbitrarily dividing bidders with attribute a into pairs and creating a new unique attribute for each pair. Then, the SS2 mechanism runs the Single Sample mechanism on Env'.

Theorem 3.16 *For every regular matroid environment with an even number of bidders of every present attribute, the expected revenue of the Single Sample 2 mechanism is at least a $\frac{1}{16}$ fraction of that of an optimal mechanism for the environment.*

The assumption that there is an even number of bidders of each attribute can be removed at the expense of a worse constant approximation factor. The high-level idea of the proof is to combine arguments used in the proof of Lemma 3.6 with a non-trivial reduction to a "Bulow-Klemperer-type" result in [11].

3.5 Computationally Efficient Variants

In the second step of the Single Sample mechanism, a different mechanism can be swapped in for the VCG mechanism. One motivation for using a different mechanism is computational efficiency (although this is not the chief focus of this paper). For example, for combinatorial auctions with single-minded bidders, implementing the VCG mechanism requires the solution of a packing problem that is NP-hard, even to approximate.

By inspection, the proof of Theorem 3.8 implies the following more general statement: if step (2) of the Single Sample mechanism uses a truthful mechanism guaranteed to produce a solution with at least a $1/c$ fraction of the maximum welfare, then the expected revenue of the corresponding Single Sample mechanism is at least a $\frac{1}{4c}\frac{\kappa-1}{\kappa}$ fraction of the expected optimal welfare (whatever the underlying m.h.r. environment). For example, for knapsack auctions — where each bidder has a public size and feasible sets of bidders are those with total size at most a given amount — we can substitute the truthful FPTAS of Briest et al. [2]. For combinatorial auctions with single-minded bidders, we can use the algorithm of Lehmann et al. [14] to obtain an $O(\sqrt{m})$-approximation in polynomial time. This factor is optimal for polynomial-time approximation (up to constant factors), under suitable complexity assumptions [14].

4. REVENUE GUARANTEES WITH MULTIPLE SAMPLES

Increasing the number of samples from the underlying valuation distributions should allow for better performance. This section modifies the Single Sample mechanism to achieve such improved guarantees, and provides quantitative and distribution-independent polynomial bounds on the number of samples required to achieve a given approximation factor.

4.1 Estimating Monopoly Reserve Prices

Improving the revenue guarantees of Section 3 via multiple samples requires thoroughly understanding the following simpler problem: Given an accuracy parameter ϵ and a regular distribution F, how many samples m from F are needed to compute a reserve price r that is $(1-\epsilon)$-*optimal*, meaning that $\hat{R}(r) \geq (1-\epsilon)\cdot\hat{R}(r^*)$ for a monopoly reserve price r^* for F? (Recall from Section 3.2 that $\hat{R}(p) = p\cdot(1-F(p))$.) We pursue bounds on m that depend *only* on ϵ and not on the distribution F — such bounds do not follow from the Law of Large Numbers and must make use of the regularity assumption.

Given m samples from F, $v_1 \geq v_2 \geq \cdots \geq v_m$, an obvious idea is to use the reserve price that is optimal for the corresponding empirical distribution, which we call the *empirical reserve*:

$$\operatorname*{argmax}_{i\geq 1} i\cdot v_i. \tag{9}$$

Interestingly, this naive approach does *not* in general give distribution-independent polynomial sample complexity bounds. Precisely, we show in the full version that for every m there is a regular distribution F such that, with constant probability, the empirical reserve (9) is not $\frac{1}{2}$-optimal. Our solution is to forbid the largest samples from acting as reserve prices, leading to a quantity we call the *guarded empirical reserve* (w.r.t. an accuracy parameter ϵ):

$$\operatorname*{argmax}_{i\geq \epsilon m} i\cdot v_i. \tag{10}$$

We can use the guarded empirical reserve to prove distribution-independent polynomial bounds on the sample complexity needed to estimate the monopoly reserve of a regular distribution.

Lemma 4.1 (Estimating the Monopoly Reserve) *For every regular distribution F and sufficiently small $\epsilon, \delta > 0$, the following statement holds: with probability at least $1-\delta$, the guarded empirical reserve (10) of $m \geq c(\epsilon^{-3}(\ln\epsilon^{-1} + \ln\delta^{-1}))$ samples from F is a $(1-\epsilon)$-optimal reserve, where c is a constant that is independent of F.*

PROOF. Set $\gamma = \epsilon/11$ and consider m samples $v_1 \geq v_2 \geq \cdots \geq v_m$ from F. Define $q_t = 1 - F(v_t)$ and $q^* = 1 - F(r^*)$, where r^* is a monopoly price for F. Since the q's are i.i.d. samples from the uniform distribution on $[0,1]$, the expected value of the quantile q_t is $t/(m+1)$, which we estimate by t/m for simplicity. An obvious approach is to use Chernoff bounds to argue that each q_t is close to this expectation, followed by a union bound. Two issues are: for small t's, the probability that t/m is a very good estimate of q_t is small; and applying the union bound to such a large number of events leads to poor probability bounds. In the following, we restrict attention to a carefully chosen small subset of quantiles, and take advantage of the properties of the normalized revenue functions of regular distributions to get around these issues.

First we choose an integer index sequence $0 = t_0 < t_1 < \cdots < t_L = m$ in the following way. Let $t_0 = 0$ and $t_1 = \lfloor \gamma m \rfloor$. Inductively, if t_i is defined for $i \geq 1$ and $t_i < m$, define t_{i+1} to be the largest integer in $\{1,\ldots,m\}$ such that $t_i < t_{i+1} \leq (1+\gamma)t_i$. If $m = \Omega(\gamma^{-2})$, then $t_i + 1 \leq (1+\gamma)t_i$ for every $t_i \geq \gamma m$ and hence such a t_{i+1} exists. Observe that $L \approx \log_{1+\gamma}\frac{1}{\gamma} = O(\gamma^{-2})$ and $t_{i+1} - t_i \leq \gamma m$ for every $i \in \{0,\ldots,L-1\}$.

We claim that, with probability 1, a sampled quantile q_t with $t \geq \gamma m$ differs from t/m by more than a $(1 \pm 3\gamma)$ factor only if some quantile q_{t_i} with $i \in \{1,2,\ldots,L\}$ differs from t_i/m by more than a $(1\pm\gamma)$ factor. For example, suppose that $q_t > \frac{(1+3\gamma)t}{m}$ with $t \geq \gamma m$; the other case is symmetric. Let $i \in \{1, 2, \ldots, L\}$ be such that $t_i \le t \le t_{i+1}$. Then

$$\begin{aligned} q_{t_{i+1}} &\ge q_t > \frac{(1+3\gamma)t}{m} \ge \frac{(1+3\gamma)t_i}{m} \\ &\ge \frac{(1+3\gamma)t_{i+1}}{(1+\gamma)m} \ge \frac{(1+\gamma)t_{i+1}}{m}, \end{aligned}$$

as claimed.

We next claim that the probability that q_{t_i} differs from t_i/m by more than a $(1\pm\gamma)$ factor for some $i \in \{1, 2, \ldots, L\}$ is at most $2Le^{-\gamma^3 m/4}$. Fix $i \in \{1, 2, \ldots, L\}$. Note that $q_{t_i} > (1+\gamma)\frac{t_i}{m}$ only if less than t_i samples have q-values at most $(1+\gamma)\frac{t_i}{m}$. Since the expected number of such samples is $(1+\gamma)t_i$, Chernoff bounds (e.g. [15]) imply that the probability that $q_{t_i} > (1+\gamma)\frac{t_i}{m}$ is at most $\exp\{-\gamma^2 t_i/3(1+\gamma)\} \le \exp\{-\gamma^2 t_i/4\} \le \exp\{-\gamma^3 m/4\}$, where the inequalities use that γ is at most a sufficiently small constant and that $t_i \ge \gamma m$ for $i \ge 1$. A similar argument shows that the probability that $q_{t_i} < (1-\gamma)\frac{t_i}{m}$ is at most $\exp\{-\gamma^3 m/4\}$, and a union bound completes the proof of the claim. Observe that if $m = \Omega(\gamma^{-3}(logL + \log\delta^{-1})) = \Omega(\epsilon^{-3}(\log\epsilon^{-1} + \log\delta^{-1}))$, then this probability is at most δ.

Now condition on the event that every quantile q_{t_i} with $i \in \{1, 2, \ldots, L\}$ differs from t_i/m by at most a $(1\pm\gamma)$ factor, and hence every quantile q_t with $t \ge \gamma m$ differs from t/m by at most a $(1\pm3\gamma)$ factor. We next show that there is a candidate for the guarded empirical reserve (10) which, if chosen, has good expected revenue. Choose $i \in \{0, 1, \ldots, L-1\}$ so that $t_i/m \le q^* \le t_{i+1}/m$. Define t^* as t_i if $q^* \ge 1/2$ and t_{i+1} otherwise. Assume for the moment that $q^* \le 1/2$. By the concavity of normalized revenue function $R(q)$ — recall Section 3.2 — $R(q_{t_{i+1}})$ lies above the line segment between $R(q^*)$ and $R(1)$. Since $R(1) = 0$, this translates to

$$\begin{aligned} R(q_{t^*}) \ge R(q^*) \cdot \frac{1-q_{t_{i+1}}}{1-q^*} &\ge R(q^*) \cdot \frac{1-(1+3\gamma)\left(\frac{t_i}{m}+\gamma\right)}{1-\frac{t_i}{m}} \\ &\ge (1-5\gamma) \cdot R(q^*), \end{aligned}$$

where in the final inequality we use that $\frac{t_i}{m} \le \frac{1}{2}$ and γ is sufficiently small. For the case when $q^* \ge \frac{1}{2}$, a symmetric argument (using $R(0)$ instead of $R(1)$ and q_{t_i} instead of $q_{t_{i+1}}$) proves that $R(q_{t^*}) \ge (1-5\gamma) \cdot R(q^*)$.

Finally, we show that the guarded empirical reserve also has good expected revenue. Let the maximum in (10) correspond to the index $\hat{t}$. Since $\hat{t}$ was chosen over t^*, $\hat{t} \cdot v_{\hat{t}} \ge t^* \cdot v_{t^*}$. Using that each of $q_{\hat{t}}, q_{t^*}$ is approximated up to a $(1 \pm 3\gamma)$ factor by $\hat{t}/m, t^*/m$ yields

$$\begin{aligned} R(q_t) &= q_{\hat{t}} v_{\hat{t}} \\ &\ge \frac{(1-3\gamma)\hat{t}}{m} v_{\hat{t}} \\ &\ge \frac{(1-3\gamma)t^*}{m} v_{t^*} \\ &\ge \frac{1-3\gamma}{1+3\gamma} q_{t^*} v_{t^*} \\ &= \frac{1-3\gamma}{1+3\gamma} R(q_{t^*}) \\ &\ge \frac{(1-5\gamma)(1-3\gamma)}{1+3\gamma} R(q^*) \\ &\ge (1-11\gamma)R(q^*). \end{aligned}$$

Since $\gamma = \epsilon/11$, the proof is complete. □

Remark 4.2 (Optimization for M.H.R. Distributions) The proof of Lemma 4.1 simplifies slightly and gives a better bound for m.h.r. distributions. The reason is a simple fact, first noted in [9, Lemma 4.1], that the selling probability q^* at the monopoly reserve r^* for an m.h.r. distribution is at least $1/e$. This means one can take the parameter t_1 in the proof of Lemma 4.1 to be $\lfloor m/e \rfloor$ instead of $\lfloor \gamma m \rfloor$ without affecting the rest of the proof. This saves a γ factor in the exponent of the bound on the probability that some q_{t_i} is not well approximated by t_i/m, which translates to a new sample complexity bound of $m \ge c(\epsilon^{-2}(\ln\epsilon^{-1} + \ln\delta^{-1}))$, where c is some constant that is independent of the underlying distribution. Also, this bound remains valid even for the empirical reserve (9) — the guarded version in (10) is not necessary.

4.2 The Many Samples Mechanism

In the following *Many Sample mechanism*, we assume that an accuracy parameter ϵ is given, and use m to denote the sample complexity bound of Lemma 4.1 (for regular valuation distributions) or of Remark 4.2 (for m.h.r. distributions) corresponding to the accuracy parameter $\frac{\epsilon}{3}$ and failure probability $\frac{\epsilon}{3}$. The mechanism is only defined if every present attribute is shared by more than m bidders.

(1) For each represented attribute a, pick a subset S_a of m *reserve bidders* with attribute a uniformly at random from all such bidders.

(2) Run the VCG mechanism on the sub-environment induced by the non-reserve bidders to obtain a preliminary winning set P.

(3) For each bidder $i \in P$ with attribute a, place i in the final winning set W if and only if v_i is at least the guarded empirical reserve r_a of the samples in S_a. Charge every winner $i \in W$ with attribute a the maximum of its VCG payment computed in step (2) and the reserve price r_a.

We prove the following guarantees for this mechanism.

Theorem 4.3 (Guarantees for Many Samples) *The expected revenue of the Many Samples mechanism is at least:*

(a) a $(1-\epsilon)$ fraction of that of an optimal mechanism in every i.i.d. regular matroid environment with at least $n \ge 3m/\epsilon = \Theta(\epsilon^{-4}\log\epsilon^{-1})$ bidders;

(b) a $(1-\epsilon)\frac{1}{e}$ fraction of the optimal expected welfare in every downward-closed m.h.r. environment with at least $\kappa \ge 3m/\epsilon = \Theta(\epsilon^{-3}\log\epsilon^{-1})$ bidders of every present attribute.

Bidders with i.i.d. and exponentially distributed valuations show that part (b) of the theorem is asymptotically optimal (as is part (a), obviously).

Proof. The lower bound on the number of bidders of each attribute implies that at most an $\epsilon/3$ fraction of all bidders are designated as reserve bidders. Lemmas 3.5 and 3.10 imply that the expectation, over the choice of reserve bidders, of the expected revenue of an optimal mechanism for and the expected welfare of the subenvironment induced by the non-reserve bidders is at least a $(1-\frac{\epsilon}{3})$ fraction of that in the full environment.

Now condition on the reserve bidders, but not on their valuations. Fix a non-reserve bidder i, and condition on the valuations of all other non-reserve bidders. Let t denote the corresponding VCG threshold for i and r^* a monopoly price for the valuation distribution F of i. Recall from Corollary 3.4 that, in an i.i.d. regular matroid environment, the conditional expected revenue obtained from i using the price $\max\{r^*, t\}$ is precisely what is obtained by the optimal mechanism for the subenvironment. Recall from Lemma 3.11 that this conditional expected revenue is at least a $1/e$ fraction of the conditional expected welfare obtained from bidder i by the VCG mechanism in the subenvironment.

The Many Samples mechanism uses the price $\max\{r, t\}$, where r is the guarded empirical reserve of the reserve bidders that share i's attribute. By Lemma 4.1 and our choice of m, r is $(1 - \frac{\epsilon}{3})$-optimal for F with probability at least $1 - \frac{\epsilon}{3}$. Concavity of the revenue function (cf., Figure 1) and an easy case analysis shows that, whenever r is $(1 - \frac{\epsilon}{3})$-optimal, the conditional expected revenue from i with the price $\max\{r, t\}$ is at least a $(1 - \frac{\epsilon}{3})$ fraction of that with the price $\max\{r^*, t\}$, for any value of t. Thus, the conditional expected revenue from i in the Many Samples mechanism is at least a $(1 - \frac{2\epsilon}{3})$ fraction of that of an optimal mechanism for the subenvironment and at least a $(1 - \frac{2\epsilon}{3})\frac{1}{e}$ fraction of the expected maximum welfare in the subenvironment. Removing the conditioning on the valuations of other non-reserve bidders; summing over the non-reserve bidders; and removing the conditioning on the choice of reserve bidders completes the proof of the theorem. $\square$

Remark 4.4 Our results in this section have interesting implications even in the special case of digital goods auctions. We note that there is no interference between different bidders in such an auction, so the general case of multiple attributes reduces to the single-attribute i.i.d. case (each attribute can be treated separately).

The Deterministic Optimal Price (DOP) digital goods auction offers each bidder i a take-it-or-leave offer equal to the empirical reserve of the other $n - 1$ bidders. The expected revenue of the DOP auction converges to that of an optimal auction as the number n of bidders goes to infinity, provided valuations are i.i.d. samples from a distribution with bounded support [6] or from a regular distribution [19]. However, the number of samples required in previous works [6, 19] to achieve a given degree of approximation depends on the underlying distribution F, and this dependence is necessary (as we show in the full version).

As an alternative, consider the variant of DOP that instead uses the guarded empirical reserve (10) of the other $n-1$ bidders to formulate a take-it-or-leave-it offer for each bidder. Our Lemma 4.1 implies a *distribution-independent* bound for this auction: provided the number of bidders is $\Omega(\epsilon^{-3} \log \epsilon^{-1})$, its expected revenue is at least a $(1 - \epsilon)$ fraction of the optimal auction.

Acknowledgments

We thank Jason Hartline for a number of helpful comments, including the observation in Remark 3.7.

5. REFERENCES

[1] A. Archer and É. Tardos. Truthful mechanisms for one-parameter agents. In *FOCS '01*, pages 482–491.

[2] Patrick Briest, Piotr Krysta, and Berthold Vöcking. Approximation techniques for utilitarian mechanism design. In *STOC '05*, pages 39–48.

[3] J. Bulow and P. Klemperer. Auctions versus negotiations. *American Economic Review*, 86(1):180–194, 1996.

[4] N. Devanur and J. D. Hartline. Limited and online supply and the Bayesian foundations of prior-free mechanism design. In *EC '09*, pages 41–50.

[5] S. Dughmi, T. Roughgarden, and M. Sundararajan. Revenue submodularity. In *EC '09*, pages 243–252.

[6] A. V. Goldberg, J. D. Hartline, A. Karlin, M. Saks, and A. Wright. Competitive auctions. *Games and Economic Behavior*, 55(2):242–269, 2006.

[7] A. V. Goldberg, J. D. Hartline, and A. Wright. Competitive auctions and digital goods. Technical Report STAR-TR-99.09.01, STAR Laboratory, InterTrust Tech. Corp., Santa Clara, CA, 1999.

[8] J. Hartline and A. Karlin. Profit maximization in mechanism design. In N. Nisan, T. Roughgarden, É. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*, chapter 13, pages 331–362. Cambridge University Press, 2007.

[9] J. D. Hartline, V. S. Mirrokni, and M. Sundararajan. Optimal marketing strategies over social networks. In *WWW '08*, pages 189–198.

[10] J. D. Hartline and T. Roughgarden. Optimal mechanism design and money burning. In *STOC '08*, pages 75–84.

[11] J. D. Hartline and T. Roughgarden. Simple versus optimal mechanisms. In *EC '09*, pages 225–234.

[12] J. D. Hartline and Q. Yan. Envy, truth, and revenue. In preparation, 2010.

[13] John O. Ledyard. Optimal combinatoric auctions with single-minded bidders. In *EC '07*, pages 237–242.

[14] D. Lehmann, L. I. O'Callaghan, and Y. Shoham. Truth revelation in approximately efficient combinatorial auctions. *Journal of the ACM*, 49(5):577–602, 2002.

[15] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[16] R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[17] Z. Neeman. The effectiveness of English auctions. *Games and Economic Behavior*, 43(2):214–238, 2003.

[18] J. G. Oxley. *Matroid Theory*. Oxford, 1992.

[19] I. Segal. Optimal pricing mechanisms with unknown demand. *American Economic Review*, 93(3):509–529, 2003.

[20] Robert B. Wilson. Game-theoretic approaches to trading processes. In ed T. Bewley, editor, *Advances in economic theory: Fifth world congress*. Cambridge, 1987.