Optimal Auctions for Correlated Buyers with Sampling *

Hu Fu
Microsoft Research
New England Lab
hufu@microsoft.com

Nima Haghpanah
Northwestern University
Department of Eletrical Engineering
and Computer Science
nima.haghpanah@gmail.com

Jason Hartline
Northwestern University
Department of Electrical Engineering
and Computer Science
hartline@eecs.northwestern.edu

Robert Kleinberg
Cornell University
Department of Computer Science
rdk@cs.cornell.edu

June 5, 2014

Abstract

Crémer and McLean [1985] showed that, when buyers' valuations are drawn from a correlated distribution, an auction with full knowledge on the distribution can extract the full social surplus. We study whether this phenomenon persists when the auctioneer has only incomplete knowledge of the distribution, represented by a finite family of candidate distributions, and has sample access to the real distribution. We show that the naive approach which uses samples to distinguish candidate distributions may fail, whereas an extended version of the Crémer-McLean auction simultaneously extracts full social surplus under each candidate distribution. With an algebraic argument, we give a tight bound on the number of samples needed by this auction, which is the difference between the number of candidate distributions and the dimension of the linear space they span.

*N. Haghpanah and J. Hartline are supported by NSF CAREER Award CCF-0846113 and NSF CCF-1101717. N. Haghpanah is also supported by a Simons Graduate Award 285006. R. Kleinberg acknowledges support from NSF award AF-0910940, a Microsoft Research New Faculty Fellowship, and a Google Research Grant.

1 Introduction

Revenue maximization, a.k.a. optimal auction design, is one of the most studied topic in the literature of mechanism design. The foundational work of Myerson [14] gave the optimal auction for selling a single item when the participating agents' values for the item are each drawn independently from distributions that are known to the auctioneer. As noted by Myerson himself, the independence of the values is crucial for his auction's optimality. Crémer and McLean [6] showed that, for correlated buyers, with a mild condition on the joint distribution, there is in general a dominant strategy incentive compatible auction that extracts full social surplus. This means that, the item is always sold to the bidder with the highest value, but the utility of each bidder is zero in expectation, and all the value created by the auction is extracted by the payments made to the auctioneer.

Both Myerson's and Crémer and McLean's auctions are subject to the criticism of the "Wilson's principle" [20], which proposes that auction design should rely as little as possible on the details of the valuation distributions, since these distributional prior knowledge can be expensive, if not impossible, to acquire. Recently, several *prior-independent* auctions were studied [e.g. 8, 19, 10]. These auctions greatly reduce the amount or precision of the knowledge needed by the auctioneer on the prior distribution, while guaranteeing nearly optimal performance against the fully knowledgeable optimal auction. However, they were all designed for independent value settings. In this work, we pursue prior-independent auction design for correlated buyers.

Typically, a prior-independent auction is an auction that is guaranteed to have nearly optimal performance on a family of distributions. To be more precise, for each distribution of a given family, the auction extracts nearly as much revenue as the optimal auction designed specifically for that distribution. For example, Dhangwatnotai et al. [8] showed that, with one sample from the underlying distribution for each bidder, a single auction can extract a revenue nearly optimal, and the family of distributions considered is that of regular distributions.

Can one use a similar method to emulate Crémer and McLean [6]'s auction for correlated settings, perhaps using samples to weaken the requirement on one's precise knowledge on the value distributions? We give an affirmative answer to this question.

Theorem 1. *Given any finite family $\mathcal{D}$ of distributions satisfying the condition of Crémer and McLean [6], a single auction with sample access to the underlying distribution can extract full social surplus under each distribution in $\mathcal{D}$, where the number of samples needed is one plus $|\mathcal{D}|$ (the number of distributions in $\mathcal{D}$), minus the dimension of the linear space spanned by the distributions in $\mathcal{D}$.*

For example, when all distributions in the family are linearly independent, the auction needs only one sample. We also show that the number of samples needed is tight (Proposition 3).

Main Difficulties. Unlike the independent value settings, for correlated bidders, the optimal auctions that guarantee bidders non-negative utilities only in expectation (i.e., interim individually rational) may extract a revenue much higher than any auction which guarantee non-negative utilities with probability one (i.e., ex post individually rational). In general, the social surplus can be extracted only with the former looser constraint on the auction. However, the expectation of a bidder's utility here is taken over the distribution of the other bidders' values, on which the auctioneer has no accurate knowledge. Therefore, the disadvantage imposed by the inaccuracy of the auctioneer's knowledge is considerably greater here than in the independent value settings. We allow ourselves a finite family of distributions to partly compensate for this disadvantage.

Since we consider a finite family of distributions, a naive way to make use of the sample access is trying to distinguish the distributions by what one sees in the samples, and then running the auction optimal for the distribution which most likely produces the samples. In Section 3, we show that this approach can be very inefficient in its use of the samples.

Theorem 2. *For any positive integer k, there exist two distributions for two bidders which satisfy the Crémer-McLean condition and which, with high probabilities, cannot be distinguished by k samples. Moreover, no single auction extracts any constant fraction of the optimal revenue under both distributions.*

In contrast, the auction in Theorem 1 needs only one sample for two such distributions.

Moreover, even when the auctioneer has some confidence on his guess of the underlying distribution and runs an auction that is interim invidually rational for that distribution, the auction may run into trouble in the event that his guess is wrong (such an event is particularly likely if the only source of confidence is the samples). This is because the auction may turn out to be not invidiually rational on the actual distribution, and participants may see negative utilities in expectation. In this case, one will have to make additional assumptions on bidders' behavior in such scenarios for any meaningful analysis. Our approach avoids this problem: as long as the actual distribution is in the family $\mathcal{D}$, our auction will be individually rational.

Our Techniques. Our auction is an extension of the auction by Crémer and McLean [6] (called the CM-auction in the sequel). That auction first runs a second price auction on the reported values, and then charges each bidder a payment (or pays her a reward) which is determined solely by the other bidders' bids. Since this payment does not depend on the bidder's own strategy, it does not alter the incentive structure of the auction. These payments can be seen as the outcome of a lottery, whose randomness comes only from the other bidders' values. The lotteries are set up so that each bidder, conditional on any of her own values, makes an expected payment in the lottery that is equal to her expected utility in the second price auction. In the independent value settings, this is not possible, because the outcome of the lottery does not depend on the bidder's own value, whereas her expected utility in the second price auction does. In the case of correlated values, this becomes possible if there is enough "richness" or "variance" in the conditional distributions of the other bidders' values as the bidder's own value varies. This "richness" is shown by Crémer and McLean to be the linear independence of the conditional distributions.

In our auction, we also first run a second price, and then decide for each bidder the outcome of her lottery without using her own value. The difference of our auction from the CM-auction is that this lottery outcome will depend not only on the other bidders' values but also on the samples from the distribution. Ideally, even though the bidder's expected utility in the second price auction changes with both her own value *and* the underlying distribution, we hope to orchestrate the change in the expected lottery outcome so as to match the utility change. This requires linear independence of all the distributions over the other bidders' values and the samples, given each candidate distribution and the bidder's own value. The main technical difficulty of this work is to give a tight bound on the number of samples needed for this linear independence. As we show in the proof of Theorem 1, it boils down to showing a property for an object in algebraic geometry known as the Veronese variety.

Structure of the paper. In Section 3 we show the limit of the naive approach by giving a proof for Theorem 2. In Section 4 we prove Theorem 1, first describing our auction and then showing the bound on the number of samples for its revenue guarantee.

1.1 Other Related Works

In correlated value settings with a known distribution, Crémer and McLean [6] gives a dominant strategy incentive compatible, interim individually rational auction that extracts full social surplus, under a certain condition (1) on the distribution. Our work is an extension of this auction. The CM auction was extended by McAfee and Reny [13] and Rahman [16] to continuous type spaces.

Another line of work studies the optimal auction for the same setting (with known prior distributions), but under the stronger constraint of ex post individual rationality. Papadimitiou and Pierrakos [15] showed that calculating the optimal deterministic auction under this requirement is NP-hard, whereas Dobzinski et al. [9] showed that the optimal randomized auction can be computed in time polynomial in the size of the distribution. Ronen [17] developed a 2-approximation for the optimal revenue where the computation cost does not grow with the number of bidders, and this approach was extended by Dobzinski et al. [9] and Chen et al. [5] for better approximation ratios. These auctions are particularly simple in form, and we will use the auction by Ronen in our proof of Theorem 2. For the more general matroid settings, Roughgarden and Talgam-Cohen [18] characterized the optimal auction under various assumptions on the distribution, and Li [12] showed that a generalized VCG auction with conditional monopoly reserve prices gives e-approximation to the optimal revenue for distributions having a correlated version of monotone hazard rate.

There have been various studies on prior-independent revenue maximization [e.g. 8, 7, 19, 10], although they all assume independent value distributions. The most relevant to this work is Dhangwatnotai et al. [8]'s single-sampling auction, which showed that with one sample from each bidder's valuation distribution, the VCG auction with the samples as reserve prices gives a 4-approximation to the optimal revenue, when the distributions are regular. As an extention, Roughgarden and Talgam-Cohen [18] gave a single-sampling mechanism for the more general interdependent value settings under various assumptions, although the benchmark is the optimal revenue under ex post individual rationality. Recently, Chawla et al. [4] gave a prior-independent mechanism optimizing a non-revenue objective, i.e., that of minimizing makespan for scheduling problems.

Online pricing [e.g. 1, 2, 3] is another setting where one has to maximize revenue but faces an unknown underlying distribution, and where one can observe values drawn from it (or partial information, e.g. by observing the buyer's decision to take or leave a certain price). The difference between this and our "batch" setting is that the observations come over time, and one needs to perform well not only in the last stage, but throughout the stages on average. Also, the buyers are typically assumed to have values (or types) drawn from the same distribution, as opposed to from a correlated distribution we consider here.

2 Preliminaries

Auctions, Incentive Compatibility and Invidual Rationality. In this paper we consider the problem of auctioning one item to n bidders whose private valuations $v_1, \dots, v_n$ are drawn from an unknown correlated distribution D. Let T_i be the support of v_i, i.e., $T_i = \{v_i \mid \exists v_{-i}, D(v_i, v_{-i}) > 0\}$. Let T be the support of D. In this work we consider only discrete distributions with finite supports.

By the revelation principle, it is without loss of generality to consider auctions of the form that solicit bidders' values and map them to an *allocation* and a *payment* for each bidder. The allocation $x_i(v_1, \ldots, v_n) \in [0, 1]$ denotes the probability with which agent i is allocated the item at the reported value profile $(v_1, \ldots, v_n)$, and the payment $p_i(v_1, \ldots, v_n)$ indicates the amount of money paid by agent i at the valuation profile. Feasibility of a single-item auction requires that $\sum_i x_i(v_1, \ldots, v_n) \leq 1$, $\forall v_1, \ldots, v_n$.

The *utility* u_i of a bidder with value v_i at an outcome x_i and p_i is $v_i x_i - p_i$. An auction is said to be *dominant strategy incentive compatible (DSIC)*, if for all i, v_i, v_i' and v_{-i},

$$v_i x_i(v_i, v_{-i}) - p_i(v_i, v_{-i}) \geq v_i x_i(v_i', v_{-i}) - p_i(v_i', v_{-i}).$$

An auction is said to be *ex post individually rational (IR)* if, for all i, v_i, v_{-i},

$$v_i x_i(v_i, v_{-i}) - p_i(v_i, v_{-i}) \geq 0.$$

An auction is said to be *interim individually rational* if, for all i and v_i,

$$\mathbf{E}_{v_{-i}}\left[v_i x_i(v_i, v_{-i}) - p_i(v_i, v_{-i}) \ \mid \ v_i\right] \geq 0,$$

where v_{-i} is drawn from the conditional distribution given v_i.

The *revenue* of an auction is the sum of expected payments it collects from all bidders. In this paper we consider maximizing revenue extractable by auctions subject to dominant strategy IC condition and ex post or interim IR condition.

Notations for Distributions. We will assume that the auctioneer is guaranteed that the valuation distribution D is from a family $\mathcal{D}$ of distributions, $\mathcal{D} = \{D^1, \cdots, D^m\}$. [1] For example, the auctioneer may have an accurate knowledge of the distribution D on the value profiles $(v_1, \ldots, v_n)$, but without knowing the mapping between the identities of the bidders in the auction and the coordinates in the valuation profile. In this case, the family $\mathcal{D}$ will consist of at most $n!$ distributions, each of which is formed by performing a permutation on the coordinates in the profiles in D.

There are multiple ways to represent a distribution. In Section 3 we represent a distribution for n bidders as an n-dimensional tensor. In particular, for two bidders, a distribution D is a $|T_1| \times |T_2|$ matrix, with the entry $D(v_1, v_2)$ denoting the probability of the occurrence of (v_1, v_2). In Section 4, we will represent a distribution by a $|T|$-dimensional vector, with $D(v_1, \ldots, v_n)$ being the probability of the occurrence of valuation profile $(v_1, \ldots, v_n)$. When $T = \prod_i T_i$, the latter is simply the vectorization of the former representation.

The probability of a valuation profile v_{-i} conditioning on bidder i's value being v_i is $D(v_{-i} \mid v_i)$. We use D_{i,v_i} to denote the conditional distribution on v_{-i} given v_i. We represent it as the $|T_{-i}|$-dimensional vector, where $D_{i,v_i,v_{-i}}$ is $D(v_{-i} \mid v_i)$.

Optimal Auctions For A Known Distribution. We will need two existing results on revenue maximization with correlated bidders, under constraints of interim IR and ex post IR, respectively.

Crémer and McLean [6] showed that, under a fairly lenient condition on the value distribution, the optimal mechanism under DSIC and interim IR can extract the full social surplus. In other words, the auction maximizes the social welfare and always allocates the item to the bidder with the highest value, whereas in expectation every bidder's utility is zero.

[1] Without loss of generality we assume these distributions have the same support.

Definition 1. A valuation distribution D is said to satisfy the *Crémer-McLean condition* if, for each bidder i, the $|T_i|$ vectors $\{D_{i,v_i}\}_{v_i \in T_i}$ are linearly independent.

Theorem 3 (6)**.** *In a single item auction where the valuation distribution satisfies the Crémer-McLean condition, there is an interim IR, DSIC auction that extracts the full social surplus.*

We will call the auction in Theorem 3 the CM auction.

Ronen [17] studied an DSIC, ex post IR *lookahead* auction that 2-approximates the optimal revenue. The auction first solicits all values, then singles out the highest bidder, and runs the optimal auction for this bidder, with the value distribution conditioning on all other bidders' values and the fact that her value is above all others'.

Theorem 4 (17)**.** *The lookahead auction is DSIC, ex post IR and extracts at least half of the optimal revenue.*

The Equal Revenue Distribution. In several examples we will make use of the following *equal revenue* distribution (truncated at h): the valuation v takes on integers between 1 and h, and the probability that $v \geq k$ is equal to $\frac{1}{k}$. The equal revenue distribution has the property that, the expectation of the value is $\Omega(\log h)$, which grows unboundedly as h grows large, but the optimal revenue one can extract from it in a single-agent setting is 1.

Kronecker Products. Notations in Section 4 will be greatly shortened by the use of Kronecker products on matrices (and vectors). The Kronecker product of matrices $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ and $B = (b_{ij}) \in \mathbb{R}^{p \times q}$ is the $mp \times nq$ block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}.$$

Kronecker products are bilinear and associative, but in general are not commutative. We will use $(\otimes A)^k$ to denote the Krocker product of k copies of A. When performing Kronecker products on an m-dimensional vector, we will treat it as an $m \times 1$ matrix. The following lemma is not hard to verify.

Lemma 5. *Consider a set of linearly independent vectors $S = \{v_1, \dots v_m\}$ and, for each $i = 1, \dots, m$, a set T_i of linearly independent vectors. The set of vectors $\cup_{i=1,\cdots,m}\{u \otimes v_i\}_{u \in T_i}$ are linearly independent. In particular, for any positive integer k, the set of m^k vectors $\{u_1 \otimes \ldots \otimes u_k\}_{u_i \in S}$, are linearly independent.*

3 The Limit of the Naive Approach

The most naive approach given sample accesses is to use the samples to distinguish distributions in the given family $\mathcal{D}$ of distributions, and then run an optimal auction for the identified distribution. However, even with the knowledge of $\mathcal{D}$, the auctioneer may still need a large number of samples to even distinguish the distributions with constant confidence of being right, let alone tailor an auction for the identified distribution. Consider the following example.

Example 1. Fix a small positive real number $\epsilon < 1$. Consider two bidders whose values are generated by the following process: two values are independently drawn from the equal revenue distribution, then with probability $1 - \epsilon$, the two values are randomly assigned to the two bidders; with probability ϵ, the higher of the two values is assigned to bidder 1, and the lower to bidder 2. Call the resulting correlated distribution D^A. Define another distribution D^B by exactly the same procedure but flipping the identity of the two bidders.

Proposition 1. *It takes $\Omega(1/\epsilon^2)$ samples to correctly distinguish D^A and D^B in Example 1 with constant probability.*

Proof. We can simulate a biased coin with D^A as follows: draw a pair of values (v_1, v_2) from D^A, and if $v_1 > v_2$, return Head; if $v_1 < v_2$, return Tail; if $v_1 = v_2$, return Head and Tail with probability $\frac{1}{2}$ each. It is not hard to see that the resulting distribution over Heads and Tails is that of an ϵ-biased coin in favor of Head. The same simulation using D^B will give a distribution of an ϵ-biased coin in favor of Tail. By standard information theoretic argument [see, e.g. Theorem 6.1 in 11], we know that it takes $\Omega(\epsilon^2)$ flips of a coin to distinguish an ϵ-biased coin in favor of Heads or Tails. Therefore one needs at least as many samples to distinguish D^A and D^B. $\square$

It is not hard to verify that D^A and D^B satisfy the Crémer-McLean condition, and one can extract full social surplus which is $\Omega(\log h)$. Now we show that, without being able to distinguish the two, one auction cannot simultaneously be interim IR and approximates the optimal revenue within a constant factor under both distributions.

Theorem 6. *There is no auction that is interim IR and gets more than $O(1)$ revenue under both distributions in Example 1 .*

Together with Proposition 1, this theorem implies Theorem 2. Before proving the theorem, we first give a characterization of dominant strategy IC mechanisms, which can be easily shown by standard arguments. We omit its proof.

Lemma 7. *Given a value distribution, any two dominant strategy IC auctions with the same allocation rule differ from each other only by a payment from each bidder i that depends only on v_{-i}.* [2]

For a fixed allocation rule, we call an auction *canonical* if it has the allocation rule, is dominant strategy IC, ex post IR and if any bidder having a value lowest in her type space always has zero utility. Given Lemma 7, we can describe any dominant strategy IC auction by the difference between it and the canonical one.

Corollary 1. *Any dominant strategy IC auction can be fully described in a standard form by its allocation rule and n vectors $L_1 \in \mathbb{R}^{|T_{-1}|}, \cdots, L_n \in \mathbb{R}^{|T_{-n}|}$. To run this auction, one first runs the canonical auction with the same allocation rule, and then charges each bidder i the amount $L_i(v_{-i})$ when the other bidders bid v_{-i}.*

Proof of Theorem 6. Fix a dominant strategy IC auction that is interim IR under both D^A and D^B in Example 1. Let $L_1 \in \mathbb{R}^{|T_2|}$ and $L_2 \in \mathbb{R}^{|T_1|}$ be the vectors of payments describing the auction's payments in addition to the canonical auction. Let $\vec{d}$ denote the equal revenue distribution truncated at h, i.e., $\vec{d} = [\frac{1}{2}, \frac{1}{6}, \cdots, \frac{1}{(h-1)h}, \frac{1}{h}]^\top$. Deviating from the rest of the paper, in this proof we

[2] We allow the payment to be negative, in which case we pay the bidder.

will use a $h \times h$ matrix D to represent a joint distribution for the two bidders, with $D(v_1, v_2)$ representing the probability of the profile (v_1, v_2). The distribution where each bidder's value is drawn independently from the equal revenue distribution is then represented by the symmetric rank-one matrix $A = \vec{d}\vec{d}^\top$. Let B be the upper triangle matrix whose diagonal elements are half of A's, and whose elements above the diagonal are identical with A. Then $A = B + B^\top$. Also, D^A is $(1-\epsilon)A + 2\epsilon B^\top = A + \epsilon(B^\top - B)$, and D^B is $A + \epsilon(B - B^\top)$.

We consider the revenue our auction gets from bidder 1 under D^A. It first gets the revenue as in the canonical auction, and then in addition, it gets $\vec{1}^\top \cdot [A + \epsilon(B^\top - B)]L_1$, where $\vec{1}^\top$ is the h-dimensional all-one row vector. Under D^B, this additional revenue is $\vec{1}^\top \cdot [A + \epsilon(B - B^\top)]L_1$. The sum of these two terms is $\vec{1}^\top \cdot AL_1$, which we show has to be small.

Let $r_1(M)$ denote the first row of a matrix M. Recall that the entries of $r_1(A + \epsilon(B^\top - B))$ correspond to the probabilities in D^A of the value profiles where $v_1 = 1$. By interim IR, we have

$$r_1(A + \epsilon(B^\top - B)) \cdot L_1 = r_1(A)L_1 + r_1(B^\top - B) \cdot L_1 \leq 1.$$

Similarly, for D^B we have

$$r_1(A)L_1 + r_1(B - B^\top) \cdot L_1 \leq 1.$$

Adding the two inequalities, we get $r_1(A)L_1 \leq 2$. Now recall that A represents the independent distribution, and $r_1(A)$ is simply $\frac{1}{2}\vec{d}^\top$. Hence $\vec{d}^\top \cdot L_1 \leq 4$. The other rows of A are also scaled copies of $\vec{d}$, and the scaling coefficients sum up to 1. Hence, $\vec{1}^\top AL_1 \leq 4$. Therefore, in addition to the canonical auction, the total sum of our auction's extra revenue from bidder 1 in the two distributions is bounded by 4. The same argument works for bidder 2 as well. In other words, our auction cannot extract substantially more revenue than the canonical auction for both distributions simultaneously. Therefore, to finish the proof, we only need to show that the canonical auction also extracts only a small revenue.

To show this, we invoke the lookahead auction. By Theorem 4, the revenue of any DSIC, ex post IR auction, including that of the canonical auction, is bounded by twice the revenue of the lookahead auction. Recall that the lookahead auction for two bidders uses the lower bidder's value to set a conditionally optimal price for the higher bidder. Given any distribution represented by a $h \times h$ matrix D, when bidder 1's value being v_1 and bidder 2's value is higher, the optimal price for bidder 2 is determined by the part of v_1-th row of D to the right of the diagonal element. If we set a price of $p_{1,v_1} \geq v_1$ in this scenario, in expectation we collect a revenue of

$$p_{1,v_1} \sum_{v \geq p_{1,v_1}} D(v_1, v).$$

We can do this for every v_1, and symmetrically for bidder 2 using the columns of D. The revenue of the lookahead auction can be expressed as

$$\max_{\substack{p_{1,1} \geq 1, \\ \cdots \\ p_{1,h} \geq h.}} \sum_{v_1=1}^{h} p_{1,v_1} \sum_{v \geq p_{1,v_1}} D(v_1, v) + \max_{\substack{p_{2,1} > 1, \\ \cdots \\ p_{2,h-1} > h-1.}} \sum_{v_2=1}^{h} p_{2,v_2} \sum_{v \geq p_{2,v_2}} D(v, v_2).$$

Now we substitute D by $D^A = A + \epsilon(B^\top - B)$. It is not hard to verify that the diagonal elements of D^A are the same as A, and for any $v_1 > v_2$, $D^A(v_1, v_2) = (1+\epsilon)A(v_1, v_2)$, and for any $v_1 < v_2$,

$D^A(v_1, v_2) = (1-\epsilon)A(v_1, v_2)$. Therefore, the revenue of the lookahead auction is upper bounded by $(1+\epsilon)$ times the following quantity:

$$\max_{\substack{p_{1,1}\geq 1,\\ \cdots\\ p_{1,h}\geq h.}} \sum_{v_1=1}^{h} p_{1,v_1} \sum_{v \geq p_{1,v_1}} A(v_1, v) + \max_{\substack{p_{2,1}>1,\\ \cdots\\ p_{2,h-1}>h-1.}} \sum_{v_2=1}^{h} p_{2,v_2} \sum_{v \geq p_{2,v_2}} A(v, v_2).$$

This quantity, however, is exactly the revenue of the lookahead auction for the independent distribution A, which in turn is known to be no more than 2 (since the revenue of Myerson's optimal auction is no more than 2; another way to see this is that for independent distributions, the optimal revenue for two bidders cannot be greater than the sum of optimal revenue extractable from each alone). This completes the proof. $\square$

4 The Power of Auctions with Internal Samples

In contrast to the limit we have shown for the naive approach, we consider an extension of the CM-auction in this section and prove Theorem 1.

Definition 2. The *CM auction with samples* works as follows:

1. Run the second price auction, which allocates the item to the highest bidder and charges her a payment equal to the second highest bid.
2. Draw k valuation profiles $\mathbf{s}_1, \ldots, \mathbf{s}_k$, each independently from the underlying distribution.
3. For each bidder, including those who do not win the item, charge her or pay her an amount of money that is a function of the other bidders bids v_{-i} *and* the samples $\mathbf{s}_1, \ldots, \mathbf{s}_k$. These functions and k, the number of samples needed, are to be specified later.

The difference between the CM auction with samples and the CM auction is the sampling procedure and the dependence of the lottery outcome on the samples. We now discuss setting up the lotteries outcomes in Step (3), and the number of samples we need. The former is an extension of the CM auction, whereas the latter involves nontrivial algebraic investigations.

4.1 Lottery Outcomes From Solving Linear Systems

The construction of the lotteries in Step (3) of 2 aims at extracting from bidder i the utility she would get in a pure second price auction, no matter what distribution we are under. This boils down to solving a linear system, as is the case in Crémer and McLean [6].

For each bidder i, we construct a vector u_i^{SPA} in $\mathbb{R}^{|T_i|\times m}$, where $u_{i,v_i,j}^{\mathrm{SPA}}$ is the expected utility of bidder i in the second price auction under distribution D^j and conditioning on that bidder i has value v_i. (Recall that m is the number of distributions in $\mathcal{D}$.) We draw k samples $\mathbf{s}_1, \ldots, \mathbf{s}_k$ from the underlying distribution, where each sample $\mathbf{s}_j$ is a profile of values $(\mathbf{s}_{j1}, \ldots, \mathbf{s}_{jn})$. We would like to decide on an amount to pay or charge bidder i given v_{-i} and $\mathbf{s}_1, \ldots, \mathbf{s}_k$. So we use a vector $L_i \in \mathbb{R}^{|T_{-i}|\times|T|^k}$ to denote these quantities, where $L_{i,v_{-i},\mathbf{s}_1,\ldots,\mathbf{s}_k}$ is the amount of money we charge or pay to bidder i, when the other bidders bid v_{-i} and when the samples are $\mathbf{s}_1, \ldots, \mathbf{s}_k$. To compute the expected payment under L, we need a distribution over the events that $v_{-i}, \mathbf{s}_1, \ldots, \mathbf{s}_k$ occur. Importantly, this distribution varies with both the underlying distribution D^j *and* bidder i's own value v_i. Therefore we have $|T_i| \cdot m$ vectors $F_i^{v_i,j}$ in $\mathbb{R}^{|T_{-i}| \times |T|^k}$, where $F_{i,v_{-i},\mathbf{s}_1,\ldots,\mathbf{s}_k}^{v_i,j}$ is the probability that the bidders other than i bid v_{-i} and that the k samples are $\mathbf{s}_1, \ldots, \mathbf{s}_k$, under the joint distribution D^j *and* conditioning on bidder i's own value being v_i. The expected payment that bidder i with value v_i makes in Step (3) under distribution D^j is then equal to $F_i^{v_i,j} \cdot L_i$.

Proposition 2. *The CM auction with samples is DSIC. In addition, if, given a family of distributions $\mathcal{D} = \{D^1, \cdots, D^m\}$, for each bidder i, the system of linear equations*

$$F_i^{v_i,j} \cdot L_i = u_i^{\text{SPA}}, \quad \forall v_i \in T_i, j \in [m] \tag{1}$$

has a solution L_i^*, then using L_i^* for bidder i in Step (3) *of 2 makes the auction interim IR and extracts full social surplus under each distribution $D^j \in \mathcal{D}$.*

Proof. The second price auction itself is DSIC, and in Step (3) of 2, the extra payment (or award) the bidder makes (or receives) is not affected by her own bid, the auction remains DSIC.

Now fixing any distribution $D^j \in \mathcal{D}$, and conditioning on bidder i having value v_i, the bidder's utility from the first two steps will be her conditional utility in a second price auction, i.e., $u_{i,v_i,j}^{\text{SPA}}$. Her extra payment in Step (3) will be in expectation equal to $F_i^{v_i,j} \cdot L_i^*$, which by definition of L_i^* is equal to $u_{i,v_i,j}^{\text{SPA}}$. This shows that the bidder has expected utility zero no matter which $D^j \in \mathcal{D}$ it is and no matter what her own value is. Therefore the auction is interim IR. As the item is always allocated to the highest bidder, the auction extracts the full social surplus. $\square$

We now investigate conditions that allow us to solve the linear systems $F_i^{v_i,j} \cdot L_i = u_i^{\text{SPA}}$. From this point on we will focus on the problem on a fixed bidder, and will drop the subscripts i. In general, there are no linear constraints governing the entries of the vector u^{SPA}, because it is calculated with both the probabilities in the distribution *and* the magnitude of the valuations. This means that, to be able to solve the linear equations, in general we need $\{F^{v_i,j}\}_{v_i,j}$ to be linearly independent.

By the independence of each sampling, we have

$$F_{v_{-i},\mathbf{s}_1,\ldots,\mathbf{s}_k}^{v_i,j} = D_{v_i}^j(v_{-i}) \cdot D^j(\mathbf{s}_1) \cdots D^j(\mathbf{s}_k),$$

(recall that $D_{v_i}^j(v_{-i})$ is the conditional probaiblity $D^j(v_{-i} \mid v_i)$), therefore

$$F^{v_i,j} = D_{v_i}^j \otimes (\otimes D^j)^k. \tag{2}$$

By the bilinearity of Kronecker products, in order to have $\{F^{v_i,j}\}_{v_i \in T_i}$ to be linearly independent even for a fixed j, we need $\{D_{v_i}^j\}_{v_i}$ to be linearly independent, which amounts to the Crémer-McLean condition (1) on D^j.

From this point on we will assume that each $D^j \in \mathcal{D}$ satisfies the Crémer-McLean condition, and we look at the number of samples needed to make $\{F^{v_i,j}\}_{v_i,j}$ linearly independent.

4.2 Upper Bounds on the Number of Samples Needed

We next show the main theorem in this section.

Theorem 8. *If each distribution $D^j \in \mathcal{D}$ satisfies the Crémer-McLean condition, and if the m vectors $\{D^j\}$ spans a linear space of dimension d, then with $k = m - d + 1$ samples, the set of vectors $\{F_i^{v_i,j}\}_{v_i,j}$ are linearly independent, for each bidder i.*

With Proposition 2, we immediately have the following corollary.

Corollary 2. *Under the condition in Theorem 8, the CM auction with $k = m - d + 1$ samples is DSIC, interim IR, and extracts full social surplus under each distribution $D^j \in \mathcal{D}$.*

Proof of Theorem 8. By (2) and Lemma 5, as we have the Crémer-McLean condition, it suffices to show that the m vectors $\{(\otimes D^j)^k\}_j$ are linearly independent.

Let $\{B_1, \cdots, B_d\}$ be a basis of the linear space spanned by $\{D^j\}_j$. Then for each j, we can write D^j as a linear sum of these vectors: $D^j = \sum_{\ell=1}^{d} \alpha_{j\ell} B_\ell$. Since each D^j is a distribution, its entries sum to one. Therefore, no two α_j and $\alpha_{j'}$ are scalar copies of each other, i.e., there are no $j \neq j'$ such that $\alpha_{j\ell} = \zeta \alpha_{j'\ell}$ for each ℓ, for some ζ.

We consider the Kronecker product $(\otimes D^j)^k$. By bilinearity,

$$
\begin{aligned}
\left(\otimes D^j\right)^k &= \left(\otimes \sum_{\ell=1}^{d} \alpha_{j\ell} B_\ell\right)^k \\
&= \sum_{\substack{\ell_1 + \cdots + \ell_d = k,\\ \ell_1, \cdots, \ell_d \geq 0}} \alpha_{j1}^{\ell_1} \alpha_{j2}^{\ell_2} \ldots \alpha_{jd}^{\ell_d} C_{\ell_1, \cdots, \ell_d},
\end{aligned}
$$

where $C_{\ell_1, \cdots, \ell_d}$ is the sum of terms that are Kronecker products of $B_1, \cdots, B_d$, such that in each term B_1 appears ℓ_1 times, and so on. (since taking kronecker product is not commutative, these products do not have to be the same.) for example, when d is two, $C_{1,2} = B_1 \otimes B_2 \otimes B_2 + B_2 \otimes B_1 \otimes B_2 + B_2 \otimes B_2 \otimes B_1$. by lemma 5, the set of vectors $\{B_{\ell_1} \otimes \cdots \otimes B_{\ell_k}\}_{\ell_1, \cdots, \ell_k \in [d]}$ are linearly independent, and therefore so are the $C_{\ell_1, \ldots, \ell_d}$'s.

now each $(\otimes D^j)^k$ is expressed as a linear combination of linearly independent vectors, with the linear coefficient on $C_{\ell_1, \cdots, \ell_d}$ being the product $\alpha_{j1}^{\ell_1} \ldots \alpha_{jd}^{\ell_d}$. to show linear independence of the set of vectors $\{(\otimes D^j)^k\}_j$, we only need to show that the set of m linear coefficients as vectors are linearly independent.

the vector $(\alpha_{j1}^{\ell_1} \ldots \alpha_{jd}^{\ell_d})_{\ell_1 + \cdots + \ell_d = k}$ is the image of the vector $\vec{\alpha}_j = (\alpha_{j1}, \ldots, \alpha_{jd})$ under the mapping $\nu : \diagdown^d \to \diagdown^{\binom{d+k-1}{d-1}}$ which evaluates all the k-th degree monomials in $\diagdown[x_1, \ldots, x_d]$ at a point in $\diagdown^d$. we now show that these m images $\nu(\vec{\alpha}_1), \ldots, \nu(\vec{\alpha}_m)$ are linearly independent when $k = m - d + 1$.

we will show that for every j, there exists a linear form on $\diagdown^{\binom{d+k-1}{d-1}}$ that vanishes at $\nu(\vec{\alpha}_{j'})$ for all $j' \neq j$ and does not vanish at $\nu(\vec{\alpha}_j)$. this will show that there cannot be any linear dependence among the m points $\nu(\vec{\alpha}_j)$.

since $\{D^j\}_j$ spans a linear space of dimension d, and since $\{B_1, \cdots, B_d\}$ is a basis of this space, the vectors $\vec{\alpha}_1, \ldots, \vec{\alpha}_m$ span a d-dimensional linear space. without loss of generality, consider $\vec{\alpha}_1$, we can find $d-1$ other vectors that are linearly independent with $\vec{\alpha}_1$. therefore we can find a linear form $f_1 : (y_1, \ldots, y_d) \mapsto \beta_1 y_1 + \cdots + \beta_d y_d$ which vanishes at all these $d-1$ vectors but does not vanish at $\vec{\alpha}_j$. without loss of generality, let the remaining $m-d$ vectors be $\vec{\alpha}_{d+1}, \ldots, \vec{\alpha}_m$. for each $j' = d+1, \cdots, m$, since $\vec{\alpha}_{j'}$ is not a scaled copy of $\vec{\alpha}_j$, we can find a linear form $f_{j'}$ such that $f_{j'}$ vanishes at $\vec{\alpha}_{j'}$ but does not vanish at $\vec{\alpha}_j$. Now consider the product of these $m-d+1$ linear forms,

$$f = f_1 f_{d+1} \ldots f_m.$$

If we take k to be $m - d + 1$, f itself is a linear form on $\mathbb{R}^{\binom{d+k-1}{d-1}}$, and can be evaluated at $\nu(\vec{\alpha}_1), \ldots, \nu(\vec{\alpha}_m)$, and

$$f(\nu(\vec{\alpha})) = f_1(\vec{\alpha}) f_{d+1}(\vec{\alpha}) \ldots f_m(\vec{\alpha}), \quad \forall \vec{\alpha} \in \mathbb{R}^d.$$

By construction, $f(\nu(\vec{\alpha}_j)) = 0$ for all $j \neq 1$ and $f(\nu(\vec{\alpha}_1)) \neq 0$. Since the choice of $\vec{\alpha}_1$ was arbitrary, the construction works for arbitrary $\vec{\alpha}_j$, and so $\nu(\vec{\alpha}_1), \ldots, \nu(\vec{\alpha}_m)$ are linearly independent for $k = m - d + 1$. This completes the proof. □

Remark 1. In the last part of the proof, since no two $\vec{\alpha}_j, \vec{\alpha}_{j'}$ are linear copies of each other, the m vectors $\vec{\alpha}_1, \ldots, \vec{\alpha}_m$ can be seen as points in the projective space $\mathbb{P}^{d-1}$. The mapping $\nu_k : \mathbb{P}^{d-1} \to \mathbb{P}^{\binom{d+k}{d}-1}$ is known as the Veronese embedding, and its image the Veronese variety. In the special case when d is two, the fact that no $k+1$ points on $\nu_k(\mathbb{P}^1)$ are linearly dependent can be somewhat more directly shown by an application of the Vandermonde determinant.

4.3 A Worst Case Lower Bound on the Number of Samples Needed

We now show that the number of samples specified in Theorem 8 is tight.

Proposition 3. *For any m and $d < m$, there exist m distributions $\{D^j\}_j$ spanning a d-dimensional linear space, such that for any $k \leq m - d$, the set of vectors $\{F^{v_i,j}\}_{v_i,j}$ are not linearly independent, for at least one bidder i.*

Proof. We first show that there are D^j's that make $\{(\otimes D^j)^k\}_j$ linearly dependent for any $k \leq m-d$. First consider the case $d = 2$.

Let B_1 and B_2 be two independent vectors in the span of D^j's. Then each D^j can be written as $\alpha_{j1}B_1 + \alpha_{j2}B_2$. Following a similar calculation as in the proof of Theorem 8, we have

$$\begin{aligned} \left(\otimes D^j\right)^k &= \left(\otimes(\alpha_{j1}B_1 + \alpha_{j2}B_2)\right)^k \\ &= \sum_{\substack{\ell_1+\ell_2=k,\\ \ell_1,\ell_2 \geq 0}} \alpha_{j1}^{\ell_1} \alpha_{j2}^{\ell_2} C_{\ell_1 \ell_2}, \end{aligned}$$

where $C_{\ell_1,\cdots,\ell_d}$ is the sum of terms that are Kronecker products of B_1 and B_2, such that in each term B_1 appears ℓ_1 times, and B_2 appears ℓ_2 times. For example, $C_{1,2} = B_1 \otimes B_2 \otimes B_2 + B_2 \otimes B_1 \otimes B_2 + B_2 \otimes B_2 \otimes B_1$. But we have just shown that all the m vectors $(\otimes D^j)^k$ can be written as a linear combinations of $k+1$ vectors, $C_{0,k}, C_{1,k-1}, \cdots, C_{\mathbf{s},0}$. Therefore, For $k < m-1$, the vectors $(\otimes D^j)^k$ cannot be linearly independent.

Now by (2), as long as we can construct, for one v_i, such that the conditional distribution $D^j_{v_i}$ is the same for all j, then the vectors $F^{v_i,j}$ cannot be linearly independent as well. This is easy to do, since $D^j_{v_i}$ only concerns a proper subset of coordinates of D^j, and we have complete freedom to construct the rest of the distribution.

The general case $d > 2$ is an easy generalization of the case of $d = 2$: given d linearly independent $D^1, \cdots, D^d$, we can always let the remaining distributions be linear combinations of D^1 and D^2, and repeat the calculation above. □

5 Discussion

Criticism on the Auction of Crémer and McLean. The surplus-extracting auction of Crémer and McLean is often seen as a critique on the model of auction design for correlated agents. The (arguably) counterintuitive phenomenon of surplus extraction is "blamed" on the unrealistic combination of several assumptions in the model: first, that the agents are risk neutral and only considers their expected linear utilities; second, that the auctioneer has exact knowledge on the underlying distribution for the agents' values; and third, that the agents themselves have the same exact knowledge. The second assumption, and the auctioneer's heavy use of this prior knowledge, is seen as a violation of the desired Wilson's principle. Our result suggests tha the precision of the auctioneer's prior knowledge may not be the main cause of the mechanism's anomalous performance — this requirement can be weakened, as long as sampling from the underlying distribution is available, and the number of samples does not have to be large. This suggests fine-tuning criticism on these auctions on the agents' precise prior knowledge and the interim individual rationality assumption.

In general, in Bayesian mechanism design, assumptions such as "who knows what" are crucial modeling decisions. Our approach via sample complexity may be useful in examining mechanisms' sensitivity to these assumptions and hence help with fine-tuning the modeling process.

Beyond Finiteness. Even though our approach involves inverting matrices whose entries are probabilities of atom events, there may be hope to extend the approach to infinite-support distributions, since there have been such extensions to Crémer and McLean's auction [e.g. 13, 16]. This seems a prerequisite for possibly extending the approach further to infinite families of distributions. We think it would be interesting to either show an impassable gap between infinite and finite families, or give conditions that makes surplus extraction possible with finitely many samples on infinite families.

6 Acknowledgement

The authors would like to thank Nick Gravin and Gjergji Zaimi for helpful discussions.

References

[1] Babaioff, M., Blumrosen, L., Dughmi, S., and Singer, Y. (2011). Posting prices with unknown distributions. In *ICS*, pages 166–178.

[2] Babaioff, M., Dughmi, S., Kleinberg, R., and Slivkins, A. (2012). Dynamic pricing with limited supply. In *ACM Conference on Electronic Commerce*, pages 74–91.

[3] Badanidiyuru, A., Kleinberg, R., and Slivkins, A. (2013). Bandits with knapsacks. In *FOCS*, pages 207–216.

[4] Chawla, S., Hartline, J. D., Malec, D. L., and Sivan, B. (2013). Prior-independent mechanisms for scheduling. In *STOC*, pages 51–60.

[5] Chen, X., Hu, G., Lu, P., and Wang, L. (2011). On the approximation ratio of k-lookahead auction. In *WINE*, pages 61–71.

[6] Crémer, J. and McLean, R. P. (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):345–61.

[7] Devanur, N. R., Hartline, J. D., Karlin, A. R., and Nguyen, C. T. (2011). Prior-independent multi-parameter mechanism design. In *WINE*, pages 122–133.

[8] Dhangwatnotai, P., Roughgarden, T., and Yan, Q. (2010). Revenue maximization with a single sample. In *ACM Conference on Electronic Commerce*, pages 129–138.

[9] Dobzinski, S., Fu, H., and Kleinberg, R. D. (2011). Optimal auctions with correlated bidders are easy. In *STOC*, pages 129–138.

[10] Fu, H., Hartline, J. D., and Hoy, D. (2013). Prior-independent auctions for risk-averse agents. In *ACM Conference on Electronic Commerce*, pages 471–488.

[11] Karp, R. M. and Kleinberg, R. (2007). Noisy binary search and its applications. In *SODA*, pages 881–890.

[12] Li, Y. (2013). Approximation in mechanism design with interdependent values. In *ACM Conference on Electronic Commerce*, pages 675–676.

[13] McAfee, R. P. and Reny, P. J. (1992). Correlated information and mecanism design. *Econometrica*, pages 395–421.

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

[15] Papadimitiou, C. and Pierrakos, G. (2011). On optimal single-item auctions. In *STOC*, pages 119–128.

[16] Rahman, D. (2014). Surplus extraction on arbitrary type spaces. Forthcoming in Theoretical Economics.

[17] Ronen, A. (2001). On approximating optimal auctions. In *ACM Conference on Electronic Commerce*, pages 11–17.

[18] Roughgarden, T. and Talgam-Cohen, I. (2013). Optimal and near-optimal mechanism design with interdependent values. In *ACM Conference on Electronic Commerce*, pages 767–784.

[19] Roughgarden, T., Talgam-Cohen, I., and Yan, Q. (2012). Supply-limiting mechanisms. In *ACM Conference on Electronic Commerce*, pages 844–861.

[20] Wilson, R. B. (1987). Game-theoretic approaches to trading processes. In Bewley, T., editor, *Advances in economic theory: Fifth world congress*, pages 33–77. Cambridge.