Mechanism Design with Unknown Correlated Distributions: Can We Learn Optimal Mechanisms?

Michael Albert
Duke University
malbert@cs.duke.edu

Vincent Conitzer
Duke University
conitzer@cs.duke.edu

Peter Stone
University of Texas at Austin
pstone@cs.utexas.edu

ABSTRACT

Due to Cremer and McLean (1985), it is well known that in a setting where bidders' values are correlated, an auction designer can extract the full social surplus as revenue. However, this result strongly relies on the assumption of a common prior distribution between the mechanism designer and the bidders. A natural question to ask is, can a mechanism designer distinguish between a set of possible distributions, or failing that, use a finite number of samples from the true distribution to learn enough about the distribution to recover the Cremer and Mclean result? We show that if a bidder's distribution is one of a countably infinite sequence of potential distributions that converges to an independent private values distribution, then there is no mechanism that can guarantee revenue more than ϵ greater than the optimal mechanism over the independent private value mechanism, even with sampling from the true distribution. We also show that any mechanism over this infinite sequence can guarantee at most a $(|\Theta| + 1)/(2 + \epsilon)$ approximation, where $|\Theta|$ is the number of bidder types, to the revenue achievable by a mechanism where the designer knows the bidder's distribution. Finally, as a positive result, we show that for any distribution where full surplus extraction as revenue is possible, a mechanism exists that guarantees revenue arbitrarily close to full surplus for sufficiently close distributions. Intuitively, our results suggest that a high degree of correlation will be essential in the effective application of correlated mechanism design techniques to settings with uncertain distributions.

Keywords

Mechanism Design; Correlated Mechanism Design; Prior Dependent Mechanisms; Auction Theory; Game Theory

1. INTRODUCTION

Auctions are widely employed at every level of the modern economy, from an individual purchasing a used CD on eBay to a multi-national corporation acquiring offshore oil drilling rights. Most often, though not exclusively, the purpose of these auctions is to generate money for the seller. Therefore, much of the work in mechanism design has been on designing mechanisms to achieve this goal of optimal revenue.

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

Revenue optimal auctions are prior dependent mechanisms, i.e., they depend on the seller knowing the distribution of bidders that may participate. The most famous example of this is the Myerson auction [17] which in many common applications is equivalent to a second price auction with a reserve. However, to effectively implement even this common, relatively simple auction, the seller must know, with a high degree of accuracy, the distribution of bidder valuations that she[1] is likely to see. If the seller overestimates the likelihood of high valuation bidders, she may set the reserve price too high and end up not selling the item at all. Therefore, a sophisticated and excellent literature [4, 7, 9, 14, 15, 16, 18] has developed techniques to optimally learn the prior distribution and construct mechanisms given samples from the true distribution.

However, the literature has, primarily, focused on the restrictive case of independent private value (IPV) distributions, where each bidder's valuation is independent of all other bidders. In the more general setting of correlated valuation distributions, i.e. settings where one bidder's valuation is correlated with other bidders' valuations, much less is known. Moreover, the correlated valuation setting is unique in that it allows for the strongest possible result in revenue maximizing mechanism design, that of full surplus extraction as revenue for the seller [2, 5, 6]. Essentially, with a small degree of correlation (a condition that we will refer to as the Albert-Conitzer-Lopomo (ACL) condition), the seller can, in expectation, generate as much revenue as if she knew the bidders' true valuations. Further, correlated valuation is likely to be the norm, not the exception, in mechanism design settings because any valuation model with a common value component will be correlated.

However, the optimal mechanisms over correlated valuations are rarely, if ever, seen in practice due to the requirement that the mechanism designer knows precisely the prior distribution over bidders' values [1, 3, 11]. If the mechanism designer tries to naïvely use an estimate of the distribution, the mechanism is unlikely to be incentive compatible or individually rational, leading to mechanisms that are hard to reason about and may perform very poorly. Therefore, if a mechanism designer intends to maximally exploit a correlated valuations setting, she must learn the distribution.

Settings with unknown, correlated valuation distributions have received relatively little attention in the literature, and therefore there is much that is unknown about the optimal mechanism design procedures. A recent paper [3] uses

[1] In this paper, we will use "he" to denote bidders and "she" to denote mechanism designers/sellers.

automated mechanism design techniques to construct mechanisms for correlated valuation settings that are robust to uncertainty in the distribution, and they demonstrate good performance in simulation. However, they provide no theoretical guarantees about the performance of the mechanism.

A seminal paper by Fu et. al. (2014) [7] explores the sample complexity of optimal mechanism design with correlated valuations and are able to show that if there is a finite set of distributions from which the true distribution will be drawn, then the sample complexity is of the same order as the number of possible distributions. However, the results are in a sense *too* strong. Specifically, their findings suggest that maximizing revenue from settings with correlated distributions with finite types is trivial from a sample complexity standpoint, at least if the set of possible distributions is known. Moreover, outside of a very small condition (effectively stating that there *is* correlation), the degree of correlation does not play a role in the ability to implement the mechanism, an intuitively strange result. The key to reconciling this intuition with their results is realizing that there is something fundamentally distinct between infinite sets of distributions and finite sets, and that their results do not extend to the case of infinite sets of distributions. Moreover, in any setting of practical interest, the mechanism designer will face an infinite number of potential distributions.

In this paper, we map the boundaries of learning in correlated valuation settings with infinite sets of distributions. Specifically, we first consider the case of a countably infinite sequence of distributions, each satisfying the ACL condition, that converges to an IPV distribution. We derive this negative result: *no mechanism can guarantee revenue any higher than the optimal revenue for the IPV distribution over the entire sequence.* Moreover, *this remains true for any mechanism that has access to a finite number of samples from the underlying distribution.* This implies that any mechanism that has access to a finite number of samples from the true distribution guarantees at most an approximation ratio of $(|\Theta|+1)/(2+\epsilon)$, where $|\Theta|$ is the number of possible bidder types. Finally, in contrast to our negative results, we show that if the true distribution satisfies the ACL condition, then there is a mechanism such that for any distribution "close enough" to the true distribution, the mechanism achieves nearly optimal revenue. Intuitively, our results suggest *that the degree of correlation will be essential in the effective application of mechanism design techniques to correlated valuation settings when the distribution is uncertain.*

2. PRELIMINARIES

We consider a single monopolistic seller auctioning one object, which the seller values at zero, to a single bidder whose valuation is correlated with an external signal. The special case of a single bidder and an externally verifiable signal captures many of the important aspects of the more general multi-agent problem of multiple bidders with correlated valuations while increasing ease of exposition, and this setting has been used in the literature on correlated mechanism design [1, 2, 3, 13] for this purpose. The external signal can, but does not necessarily, represent other bidders' bids.

The bidder has a type θ drawn from a finite set of discrete types $\Theta = \{1, ..., |\Theta|\}$. Further, the bidder has a valuation function $v : \Theta \rightarrow \mathbb{R}^+$ that maps types to valuations for the object. Assume without loss of generality that for all $\theta, \theta' \in \Theta$, if $\theta > \theta'$ then $v(\theta) \geq v(\theta')$. The discrete external signal is denoted by $\omega \in \Omega = \{1, 2, \ldots, |\Omega|\}$. Throughout the paper, we will denote vectors, matrices, and tensors as bold symbols, but elements of these as standard type.

There is a probability distribution, $\boldsymbol{\pi}$, over the types of the bidder and external signal where the probability of type and signal (θ, ω) is $\pi(\theta, \omega)$. The probability distribution can be represented in many possible ways, but we will represent it as a matrix. Specifically, the distribution is a matrix of dimension $|\Theta| \times |\Omega|$ whose elements are all positive and sum to one. Note that in contrast to much of the literature on mechanism design, we do not require that the bidder type be distributed independently of the external signal.

The distribution over the external signal ω given θ will be denoted by the $|\Omega|$ dimensional vector $\boldsymbol{\pi}(\cdot|\theta)$. We are, throughout this paper, primarily interested in the conditional distribution over the external signal given the bidder's type, $\boldsymbol{\pi}(\cdot|\theta)$, so we will represent the full distribution as a marginal distribution over Θ, $\boldsymbol{\pi}_\theta$, and a set of conditional distributions over Ω, $\boldsymbol{\pi}(\cdot|\cdot) = \{\boldsymbol{\pi}(\cdot|1), \boldsymbol{\pi}(\cdot|2), ..., \boldsymbol{\pi}(\cdot||\Theta|)\}$.

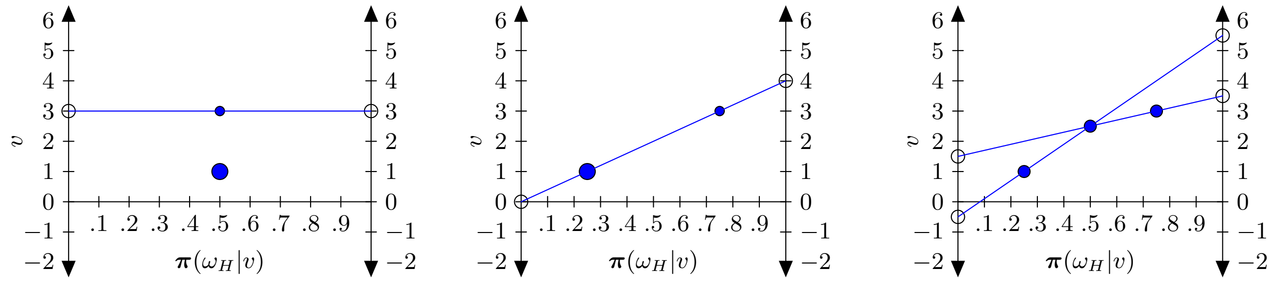
We explore the case where the bidder knows the prior distribution, but the seller does not. Specifically, throughout the paper we will assume that there is a set of distributions $\{\boldsymbol{\pi}_i\}_{i\in Q}$ where Q is some set, and only the bidder knows which distribution is the true distribution. We will assume that the seller does know the marginal distribution over θ, $\boldsymbol{\pi}_\theta$, and that the marginal distribution is fixed, i.e. all potential distributions have the same marginal distribution. We can assume, without loss of generality, that the smallest value for the marginal distribution $\min_{\theta'\in\Theta} \pi_\theta(\theta') > 0$, and we will denote it as $\pi_{\min}$ We are interested in understanding the limitations to learning specifically in the correlated valuation setting, so assuming that the marginal distribution is known and fixed allows us to explore the additional challenges of learning the conditional distribution without unnecessary complicating factors. All of our results hold if the marginal distribution is uncertain as well. Given that we are primarily interested in the conditional distributions, we will denote the full set of conditional distributions, i.e. $\cup_{\theta\in\Theta}\{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i\in Q}$, as $\{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i\in Q,\theta}$. We allow the bidder to report both his type θ and his distribution $\boldsymbol{\pi}_i$, and the mechanism may be conditional on these reports.

A (direct) revelation mechanism is defined by, given the bidder type, bidder distribution and external signal $(\theta, \boldsymbol{\pi}, \omega)$, 1) the probability that the seller allocates the item to the bidder and 2) a monetary transfer from the bidder to the seller. We will denote the probability of allocating the item to the bidder as an element of the $|\Theta| \times |\{\boldsymbol{\pi}_i\}_{i\in Q}| \times |\Omega|$ tensor $\boldsymbol{p}$. An element of the tensor $\boldsymbol{p}$ will be denoted by $p(\theta, \boldsymbol{\pi}, \omega)$, a value between zero and one. Similarly, the transfer from the bidder to the seller is denoted $\boldsymbol{x}$ and an element as $x(\theta, \boldsymbol{\pi}, \omega)$, where a positive value denotes a payment to the seller and a negative value a payment from the seller to the bidder. We will denote a mechanism as $(\boldsymbol{p}, \boldsymbol{x})$.

Definition 1 (Bidder's Utility). *Given a realization of the external signal ω, reported type $\theta' \in \Theta$, reported distribution $\boldsymbol{\pi}' \in \{\boldsymbol{\pi}_i\}_{i\in Q}$, true type $\theta \in \Theta$, and true distribution $\boldsymbol{\pi} \in \{\boldsymbol{\pi}_i\}_{i\in Q}$, the bidder's utility under mechanism $(\boldsymbol{p}, \boldsymbol{x})$ is:*

$$U(\theta, \boldsymbol{\pi}, \theta', \boldsymbol{\pi}', \omega) = v(\theta)p(\theta', \boldsymbol{\pi}', \omega) - x(\theta', \boldsymbol{\pi}', \omega)$$

Due to the well-known revelation principle (e.g. [8]), the seller can restrict her attention to incentive compatible mechanisms, i.e., mechanisms where it is always optimal for the bidder to truthfully report his valuation.



(a) Distribution where the valuation is uncorrelated with the external signal, π_1

(b) Distribution that satisfies the Cremer-McLean Condition, π_2

(c) Distribution that satisfies the ACL condition but fails Cremer-Mclean, π_3

Figure 1: The points represent the bidder type, where the position along the x-axis is the probability that the external signal is high. The relative size of the point represents the marginal probability of that bidder type. The lines represent lotteries offered in the mechanism, with the payment for the lottery if ω_H is observed being the intersection with the right vertical axis, and the payment if ω_L is observed is the intersection with the left axis. The height of the line at each point is the expected payment for that lottery. The bidder accepts a lottery if and only if the expected payment is less than or equal to his valuation (IR) and chooses the lottery with the lowest expected payment (IC). For these mechanisms, if a bidder accepts a lottery, the item is allocated with probability 1. Figure 1a shows a take it or leave it offer of 3, and only the high valuation $v = 3$ is allocated the item.

Definition 2 (Bayesian Incentive Compatibility). *A mechanism* $(\boldsymbol{p}, \boldsymbol{x})$ *is* Bayesian incentive compatible (IC) *if for all* $\theta, \theta' \in \Theta$ *and* $\boldsymbol{\pi}, \boldsymbol{\pi}' \in \{\boldsymbol{\pi}_i\}_{i \in Q}$*:*

$$\sum_{\omega \in \Omega} \pi(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega) \geq \sum_{\omega \in \Omega} \pi(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta', \boldsymbol{\pi}', \omega)$$

Bayesian incentive compatibility is a statement about the *beliefs* of the bidder over the external signal, $\boldsymbol{\pi}(\omega|\theta)$. Specifically, it allows the seller to determine payments by *lottery*. The lottery that bidder i faces can be dependent on his valuation, but the lottery itself is over the external signal. In order for the mechanism to be incentive compatible the bidder must *believe* that his expected utility is higher from the lottery he gets by reporting his valuation truthfully than by reporting any other valuation (see [2] for an in depth exploration of this point).

In addition to incentive compatibility, we are interested in mechanisms that are *individually rational*, i.e. it is rational for a bidder to participate in the mechanism.

Definition 3 (Ex-Interim Individual Rationality). *A mechanism* $(\boldsymbol{p}, \boldsymbol{x})$ *is* ex-interim individually rational (IR) *if for all* $\theta \in \Theta$ *and* $\boldsymbol{\pi} \in \{\boldsymbol{\pi}_i\}_{i \in Q}$*:*

$$\sum_{\omega \in \Omega} \pi(\omega|\theta) U(\theta, \boldsymbol{\pi}, \theta, \boldsymbol{\pi}, \omega) \geq 0$$

To illustrate the importance of prior-dependent mechanisms, it is necessary to review two important results in the literature on revenue maximization with correlated valuation distributions when the distribution is perfectly known.

Definition 4 (Cremer-McLean Condition). *The distribution over bidder types* $\boldsymbol{\pi}$*, is said to satisfy the Cremer-McLean condition if the set of beliefs associated with the bidder,* $\{\boldsymbol{\pi}(\cdot|\theta) : \theta \in \Theta\}$*, are linearly independent.*

Theorem 1 (Cremer and McLean 1985). *If the Cremer-McLean condition is satisfied by the distribution* $\boldsymbol{\pi}$*, then there exists an ex-interim IR and ex-post IC mechanism that extracts the full social surplus as revenue.*

This result states that under the apparently reasonable Cremer-McLean condition, i.e. a condition that holds with probability one for a random distribution [10], the mechanism designer can generate as much revenue in expectation as if she knew the bidder's valuation. This is a remarkable result and it can be relaxed further by the results in [2].

Theorem 2 (Albert, Conitzer, and Lopomo 2016). *A Bayesian IC and ex-interim IR mechanism can extract full social surplus as revenue if and only if there exists a* concave *function* $G : \mathbb{R}^{|\Omega|} \to \mathbb{R}$ *such that for all* $\theta \in \Theta$*,* $G(\boldsymbol{\pi}(\cdot|\theta)) = v(\theta)$*.*

We will refer to any distribution $\boldsymbol{\pi}$ that satisfies the conditions for full surplus extraction as in Theorem 2 as satisfying the Albert-Conitzer-Lopomo (ACL) condition. Example 1 and Figure 1 demonstrates the Cremer-McLean and ACL conditions, and gives example mechanisms that extract full surplus as revenue.

Example 1. *Suppose that there is a single bidder and an external signal that is correlated with the bidder's valuation. Both the bidder valuations and the external signal are binary, and we will denote the bidder valuations by* $v \in \{1, 3\}$ *and the possible values of the external signal by* $\omega \in \{\omega_L, \omega_H\}$*. Denote the distribution of the bidder's valuations and the external signal by*

$$\boldsymbol{\pi}_1(v, \omega) = \begin{bmatrix} 1/3 & 1/3 \\ 1/6 & 1/6 \end{bmatrix} \quad \boldsymbol{\pi}_2(v, \omega) = \begin{bmatrix} 1/2 & 1/6 \\ 1/12 & 1/4 \end{bmatrix}$$

where the indices are ordered such that $\boldsymbol{\pi}_2(v = 3, \omega = \omega_L) = 1/12$*. Note that the marginal distributions over* v *are identical for* $\boldsymbol{\pi}_1$ *and* $\boldsymbol{\pi}_2$*. It is clear that in* $\boldsymbol{\pi}_1$ *the bidder's valuation and the external signal are uncorrelated, implying that the optimal mechanism is a reserve price mechanism [17], shown in Figure 1a, with an expected revenue of* 1*. However,* $\boldsymbol{\pi}_2$ *satisfies the Cremer-McLean condition, and therefore, the seller can extract full surplus as revenue (the full* 5/3*), as in Figure 1b.*

Now suppose that the set of valuations is instead $v \in \{1, 2.5, 3\}$*, and the distribution over types and external signals is given by*

$$\boldsymbol{\pi}_3(v,\omega) = \begin{bmatrix} 1/4 & 1/12 \\ 1/6 & 1/6 \\ 1/12 & 1/4 \end{bmatrix}$$

It is trivial to verify that $\boldsymbol{\pi}_3$ does not satisfy the Cremer-McLean Condition. However, it does satisfy the ACL condition, and Figure 1c shows a mechanism that extracts full surplus.

3. CONVERGING SEQUENCES OF DISTRIBUTIONS

This section formalizes the infinite set of distributions that we will examine for our impossibility result, and then proves the main results. It is likely that any reasonable distribution estimation procedure will return a continuous and closed set of distributions that are consistent with the observed samples. Recent literature on designing correlated mechanisms when there is uncertainty in the distribution has estimated the discrete distribution using a Bayesian approach with an uninformative prior [3], specifically estimating a categorical distribution using Dirichlet posterior. This procedure will never rule out any distributions, so the possible sets of distributions (i.e. the distributions that have positive probability mass in the posterior of the estimation) are all distributions, an uncountably infinite, continuous, closed set.

In this section, we will consider a more restrictive set, that of countably infinite sequences of distributions that converge to a distribution. This set is more restrictive in the sense that any continuous closed set of distributions will contain an infinite number of these sequences. Since we will show that this more restrictive set is sufficient, the results will naturally extend to the more permissive set.

Definition 5. *A countably infinite sequence of distributions $\{\boldsymbol{\pi}_i\}_{i=1}^{\infty}$ is said to be* converging to the distribution $\boldsymbol{\pi}^*$, *the* convergence point, *if for all $\theta \in \Theta$ and $\epsilon > 0$, there exists a $T \in \mathbb{N}$ such that for all $i \geq T$, $||\boldsymbol{\pi}_i(\cdot|\theta) - \boldsymbol{\pi}^*(\cdot|\theta)|| < \epsilon$. I.e., for each $\theta \in \Theta$, the conditional distributions in the sequence, $\{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i=1}^{\infty}$, converge to the conditional distribution $\boldsymbol{\pi}^*(\cdot|\theta)$ in the l^2 norm.*

Note that in Definition 5, we do not explicitly assume that the elements of the sequence satisfy the ACL condition, nor do we assume that the distribution to which the sequence is converging is an IPV distribution. However, it is straightforward to construct examples of converging sequences such that every element of the sequence satisfies ACL but the limit is IPV. Figure 2a demonstrates one such set. We will make use of the following standard definition.

Definition 6 (Affine Independence). *A set of vectors $\{\boldsymbol{v}_i\}_{i=1}^{m}$ over $\mathbb{R}^n$ are affinely independent if for $\{\alpha_i\}_{i=1}^{m}$, $\sum_i \alpha_i \boldsymbol{v}_i = \mathbf{0}$ and $\sum_i \alpha_i = 0$ implies $\alpha_i = 0$ for all $i \in \{1, ..., m\}$.*

The set of distributions over Ω are the points on a $|\Omega|$-simplex where the vertices of the simplex are denoted by the set of distributions such that $\pi(\omega) = 1$ for all $\omega \in \Omega$ (see Figure 2b). Further, any set of distributions over Ω of size $|\Omega|$ that are affinely independent must span the $|\Omega|$-simplex with affine combinations. I.e., if the set $\{\boldsymbol{\pi}_i\}_{i=1}^{|\Omega|}$ is affinely independent, then for any distribution $\boldsymbol{\pi}'$ over Ω, there must exist $\{\alpha_i\}_{i=1}^{|\Omega|}$ where $\sum_i \alpha_i = 1$ and $\boldsymbol{\pi}' = \sum_i \alpha_i \boldsymbol{\pi}_i$.

We can assume, without loss of generality, that for any sequence of distributions we consider, $\{\boldsymbol{\pi}_i\}_{i=1}^{\infty}$, there must exist a subset of $\{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i,\theta}$ of size $|\Omega|$ that is affinely independent. If not, the affine combination of vectors $\{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i,\theta}$ spans a lower dimensional simplex, and we can reduce the dimensionality of Ω until an affinely independent subset exists. Note that this relies on the assumption that the bidder is risk neutral. Specifically, a risk neutral bidder is indifferent between a payment for an outcome of the external signal, $p(\theta', \boldsymbol{\pi}_{i'}, \omega)$, and a lottery over multiple values of the external signal with the same expected payoff. Therefore, if there is not a subset of $\{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i,\theta}$ of size $|\Omega|$ that is affinely independent we can always replace the true signal with a lower dimensional set of lotteries over the external signal without affecting the expected utility of the bidder.

In addition to Definition 5, we will require the following assumption.

Assumption 1. *For the sequence of distributions $\{\boldsymbol{\pi}_i\}_{i=1}^{\infty}$ converging to $\boldsymbol{\pi}^*$ and for any $\theta' \in \Theta$, there exists a subset of distributions of size $|\Omega|$ from the set $\{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i,\theta}$ that is affinely independent and the distribution $\boldsymbol{\pi}^*(\cdot|\theta')$ is a strictly convex combination of the elements of the subset. I.e., there exists $\{\alpha_k\}_{k=1}^{|\Omega|}$, $\alpha_k \in (0,1)$, and $\{\boldsymbol{\pi}_k(\cdot|\theta_k)\}_{k=1}^{|\Omega|}$ such that $\boldsymbol{\pi}^*(\cdot|\theta') = \sum_{k=1}^{|\Omega|} \alpha_k \boldsymbol{\pi}_k(\cdot|\theta_k)$.*

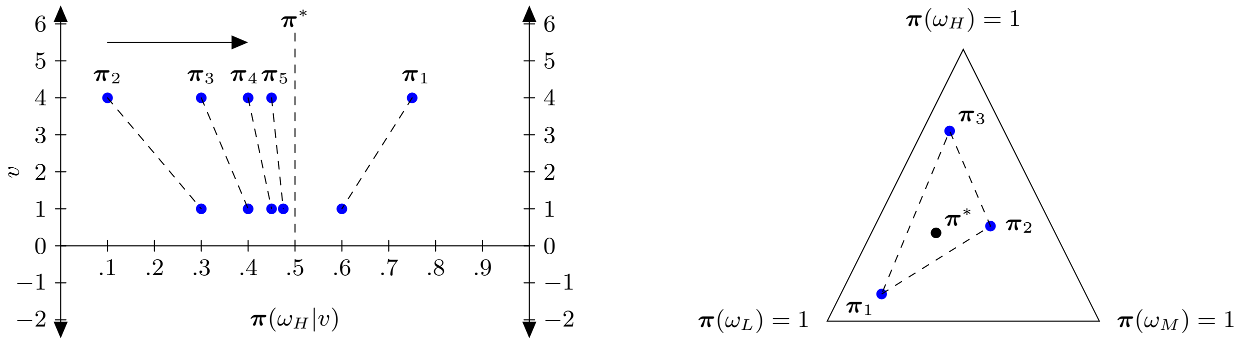
Assumption 1 states that the sequence of distributions is converging to a distribution that is in the interior of the sequence. This is not without loss of generality, but it greatly simplifies the analysis. Moreover, we believe that it is very likely to hold in realistic settings. Specifically, Assumption 1 is a statement about the *conditional distributions*, and particularly that all conditional distributions of the convergence point, $\boldsymbol{\pi}^*(\cdot|\theta)$ for all θ, is in some sense in the interior of some other estimate (see Figures 2b for a graphical depiction of this statement). However, the distributions that "enclose" the convergence point do not have to have the same θ, i.e. any conditional distributions for any θ in the set $\{\boldsymbol{\pi}_i(\cdot|\theta\}_{i,\theta}$ can be the distributions that "enclose" the convergence point.

Further, if the full set of all potential distributions is a continuous closed set and has an independent distribution in the interior of the set, then there will be an infinite number of sequences that satisfy this assumption. This is the case for the estimation procedure used in [3].

With these definitions, we are able to introduce our main results, Theorem 3 and Corollary 4.

Theorem 3. *Let $\{\boldsymbol{\pi}_i\}_{i=1}^{\infty}$ be a sequence of distributions converging to $\boldsymbol{\pi}^*$ that satisfies Assumption 1. Denote the revenue of the optimal mechanism for the distribution $\boldsymbol{\pi}^*$ by R. For any $k > 0$, and for any mechanism that is incentive compatible and individually rational, there exists a $T \in \mathbb{N}$ such that for all $\boldsymbol{\pi}_{i'} \in \{\boldsymbol{\pi}_i\}_{i=T}^{\infty}$, the expected revenue is less than $R + k$.*

Theorem 3, whose proof we shall defer to the end of this section, states that *no mechanism* can guarantee revenue better than the optimal revenue achievable at the convergence point for all distributions in the sequence. Namely, if the sequence of distributions $\{\boldsymbol{\pi}_i\}_{i=1}^{\infty}$ satisfy the ACL condition, but the convergence point is IPV, then *no mechanism*



(a) Sequence of converging distributions with a binary signal.

(b) Sequence elements that satisfy Assumption 1.

Figure 2: Figure 2a demonstrates a converging sequence as in Definition 5. Each point represents a conditional distribution, and conditional distributions linked by a dashed line both belong to the same full distribution. Specifically, the conditional distributions are all converging to $\boldsymbol{\pi}(\omega_H|v) = 1/2$, i.e. a distribution where the bidder's value and the external signal are uncorrelated. However, all of the distributions in the sequence satisfy the Cremer-McLean condition. Figure 2b demonstrates distributions that satisfy Assumption 1 in a set of distributions over three possible external signals $\{\omega_L, \omega_M, \omega_H\}$ as points in a 2-simplex. Specifically, $\boldsymbol{\pi}^*$ is a strictly convex combination of $\boldsymbol{\pi}_1$, $\boldsymbol{\pi}_2$, and $\boldsymbol{\pi}_3$. The sequence of distributions in Figure 2a also satisfies Assumption 1 due to $\boldsymbol{\pi}_1$, but if $\boldsymbol{\pi}_1$ was excluded from the sequence, it would not.

can do always do better than the optimal mechanism for the IPV point (in our setting, a reserve price mechanism [17]).

It may not seem surprising that we cannot construct mechanisms that do well on large sets of distributions. However, the following corollary indicates that *we cannot learn* a mechanism that always does well either.

COROLLARY 4. *Let $\{\boldsymbol{\pi}_i\}_{i=1}^{\infty}$ be a sequence of distributions converging to $\boldsymbol{\pi}^*$ that satisfies Assumption 1. Denote the revenue of the optimal mechanism for the distribution $\boldsymbol{\pi}^*$ by R. For any $k > 0$ and for any mechanism that is incentive compatible and individually rational and uses a finite number of independent samples from the underlying distribution, there exists a $T \in \mathbb{N}$ such that for all $\boldsymbol{\pi}_{i'} \in \{\boldsymbol{\pi}_i\}_{i=T}^{\infty}$, the expected revenue is less than $R + k$.*

It is important to be very careful in interpreting Theorem 3 and Corollary 4; they are both statements about distributions close to the convergence point. They do not provide a bound for distributions that are far from the convergence point. Therefore, even if the convergence point is an IPV distribution, it is still potentially possible to generate near optimal revenue for some distributions in the sequence, in fact we will formally show this in Theorem 10. However, even with sampling, mechanisms cannot generate significantly higher revenue than the optimal IPV mechanism for distributions sufficiently close to IPV, though sampling may still substantially increase the expected revenue for some subset of the sequence of distributions.

These results indicate that the setting where the bidder may have a distribution from an infinite set is fundamentally different from the setting where the bidder's distribution is one of a finite set (as in Fu. et. al. 2014 [7]). Note that the set of all mechanisms includes mechanisms that first applies some procedure to reduce the infinite set to a finite set.

In the remainder of this section, we prove Theorem 3 and Corollary 4. The strategy that we will use to prove the above results relies on bounding the maximum possible payments for any mechanism. Specifically, the revelation principle [8] ensures that the revenue achievable by any mechanism can be achieved by a mechanism that not only truthfully elicits the bidder's valuation, but also truthfully elicits the distribution of the bidder. We will show that Assumption 1 implies that any mechanism with payments too large (either from or to the bidder), will create an incentive for some bidder type to lie either about his valuation or his distribution, violating the revelation principle. Once we show that payments are bounded, we can use a standard continuity result in linear programming to show that the expected revenue of the mechanism must converge to something less than or equal to the optimal revenue achievable at the convergence point.

To bound payments, we will require that for distributions "sufficiently close" to the convergence point, we can always find another distribution that is a finite step in any direction. This is what Assumption 1 provides (see Figure 2b for intuition), as the following lemma formally demonstrates.

LEMMA 5. *Let $\{\boldsymbol{\pi}_i\}_{i=1}^{\infty}$ be a sequence of distributions converging to $\boldsymbol{\pi}^*$ that satisfies Assumption 1. There exists an $\epsilon_{\min} > 0$, such that for all distributions $\boldsymbol{\pi}(\cdot)$ over Ω where $||\boldsymbol{\pi}(\cdot) - \boldsymbol{\pi}^*(\cdot|\theta)|| < \epsilon_{\min}$ for some $\theta \in \Theta$, and all unit vectors $\boldsymbol{z} \in \mathbb{R}^{|\Omega|}$ where $\sum_\omega z(\omega) = 0$, there exists a $\boldsymbol{\pi}_j(\cdot|\theta_j) \in \{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i,\theta}$ such that $(\boldsymbol{\pi}(\cdot) - \boldsymbol{\pi}_j(\cdot|\theta_j)) \bullet \boldsymbol{z} \geq \epsilon_{\min}$.*

PROOF. First, note that by Assumption 1, for all $\theta \in \Theta$, there exists $\{\alpha_k\}_{k=1}^{|\Omega|}$ and an affinely independent set of vectors $\{\boldsymbol{\pi}_k(\cdot|\theta_k)\}_{k=1}^{|\Omega|}$, where $\alpha_k \in (0,1)$ and $\boldsymbol{\pi}^*(\cdot|\theta) = \sum_{k=1}^{|\Omega|} \alpha_k \boldsymbol{\pi}_k(\cdot|\theta_k)$. The set of affinely independent points $\{\boldsymbol{\pi}_k(\cdot|\theta_k)\}_k$ define a simplex in $\mathbb{R}^{|\Omega|}$, and the lth face of the simplex, where $l \in \{1, ..., |\Omega|\}$, is the set of points denoted by $\sum_{k \neq l} \alpha'_k \boldsymbol{\pi}_k(\cdot|\theta_k)$ such that $\sum_{k \neq l} \alpha'_k = 1$ and $\alpha'_k \in [0,1]$. The distance from the distribution $\boldsymbol{\pi}^*(\cdot|\theta)$ to any point on the lth face is:

$$\min_{\alpha'_k} ||\boldsymbol{\pi}^*(\cdot|\theta) - \sum_{k \neq l} \alpha'_k \boldsymbol{\pi}_k(\cdot|\theta_k)||$$

$$= \min_{\alpha'_k} ||\sum_{k=1}^{|\Omega|} \alpha_k \boldsymbol{\pi}_k(\cdot|\theta_k) - \sum_{k \neq l} \alpha'_k \boldsymbol{\pi}_k(\cdot|\theta_k)||$$

$$= \min_{\alpha'_k} ||\alpha_l \boldsymbol{\pi}_l(\cdot|\theta_l) - \sum_{k \neq l} (\alpha_k - \alpha'_k) \boldsymbol{\pi}_k(\cdot|\theta_k)|| > 0.$$

The last inequality is due to $\alpha_l \neq 0$ (Assumption 1) and affine independence. Let $\epsilon' > 0$ be the minimum such distance for all $\theta \in \Theta$ and all faces of the simplex.

Define $\epsilon_{\min} = \frac{\epsilon'}{2}$. Let $\boldsymbol{\pi}(\cdot)$ be a distribution over Ω where $||\boldsymbol{\pi}(\cdot) - \boldsymbol{\pi}^*(\cdot|\theta)|| < \epsilon_{\min}$, for some $\theta \in \Theta$. Let $\{\boldsymbol{\pi}_k(\cdot|\theta_k)\}_k$ define the simplex that contains $\boldsymbol{\pi}^*(\cdot|\theta)$. Therefore, the distance from $\boldsymbol{\pi}(\cdot)$ to any face of the simplex is at least $\epsilon_{\min}$ by an application of the triangle inequality. Let $\boldsymbol{z} \in \mathbb{R}^{|\Omega|}$ be a unit vector such that $\sum_\omega z(\omega) = 0$. Since a simplex is a closed and bounded set, there exists some face of the simplex which we will denote as the lth face, such that for some $\{\alpha'_k\}_{k \neq j}$ and some $\epsilon \geq \epsilon_{\min} > 0$:

$$\boldsymbol{\pi}(\cdot) - \epsilon \boldsymbol{z} = \sum_{k \neq l} \alpha'_k \boldsymbol{\pi}_k(\cdot|\theta_k).$$

Let $\boldsymbol{\pi}_j(\cdot|\theta_j)$ be a vertex of that face such that

$$(\boldsymbol{\pi}_j(\cdot|\theta_j) - \sum_{k \neq l} \alpha'_k \boldsymbol{\pi}_k(\cdot|\theta_k)) \bullet \boldsymbol{z} \leq 0.$$

This must exist by virtue of the face being a segment of a hyper-plane. Then:

$$\begin{aligned}
\epsilon_{\min} &\leq \epsilon = \epsilon \boldsymbol{z} \bullet \boldsymbol{z} = (\boldsymbol{\pi}(\cdot) - \sum_{k \neq l} \alpha'_k \boldsymbol{\pi}_k(\cdot|\theta_k)) \bullet \boldsymbol{z} \\
&= (\boldsymbol{\pi}_j(\cdot|\theta_j) - \sum_{k \neq l} \alpha'_k \boldsymbol{\pi}_k(\cdot|\theta_k)) \bullet \boldsymbol{z} + (\boldsymbol{\pi}(\cdot) - \boldsymbol{\pi}_j(\cdot|\theta_j)) \bullet \boldsymbol{z} \\
&\leq (\boldsymbol{\pi}(\cdot) - \boldsymbol{\pi}_j(\cdot|\theta_j)) \bullet \boldsymbol{z}
\end{aligned}$$

$\square$

As discussed in Section 2, the payments in a Bayesian mechanism are a lottery over the external signal (see Figure 1c). A lottery over the external signal can be viewed as a linear function (or a hyper-plane) whose domain is the Ω-simplex of distributions and whose value is the expected payment for the lottery. Lemma 5 ensures that for points close enough to the convergence point, there exists a distribution in the sequence that is in the "opposite direction" of the gradient of the hyperplane that defines the lottery. I.e., for any possible lottery with a gradient of magnitude K, there exists a distribution for which the expected payment for the lottery is at least $\epsilon_{\min} K$ less than for any distributions "sufficiently close" to the convergence point. Therefore, if payments are too large (either from or to the bidder) for some distribution $\boldsymbol{\pi}'$ and θ', there is another distribution and type $\boldsymbol{\pi}''$ and θ'' that will find reporting $\boldsymbol{\pi}'$ and θ' irresistable. The following lemma formalizes this argument.

LEMMA 6. *Let $\{\boldsymbol{\pi}_i\}_{i=1}^\infty$ be a sequence of distributions converging to $\boldsymbol{\pi}^*$ that satisfies Assumption 1. For any mechanism $(\boldsymbol{p}, \boldsymbol{x})$ that is incentive compatible and individually rational and guarantees non-negative revenue in expectation for all distributions in $\{\boldsymbol{\pi}_i\}_{i=1}^\infty$, there exists some $M > 0$ such that for all $\boldsymbol{\pi}_{i'} \in \{\boldsymbol{\pi}_i\}_{i=1}^\infty, \theta \in \Theta$, and $\omega \in \Omega$:*

$$|x(\theta, \boldsymbol{\pi}_{i'}, \omega)| \leq M$$

PROOF. Let $\epsilon_{\min} > 0$ be defined as in Lemma 3. By Definition 5, there exists a $T \in \mathbb{N}$ such that for all $\theta \in \Theta$ and $\boldsymbol{\pi}_{i^*}(\cdot|\theta) \in \{\boldsymbol{\pi}_i(\cdot|\theta)\}_{i=T}^\infty$, $||\boldsymbol{\pi}_{i^*}(\cdot|\theta) - \boldsymbol{\pi}^*(\cdot|\theta)|| \leq \epsilon_{\min}$. Since there are a finite number of distributions such that $i^* < T$, choose $M_{i^* < T} = \max_{i^* < T, \theta, \omega} |x(\theta, \boldsymbol{\pi}_{i^*}, \omega)|$.

Therefore if payments are not bounded, for any $M' > 0$, there must exist some $\boldsymbol{\pi}_{i'} \in \{\boldsymbol{\pi}_i\}_{i=T}^\infty, \theta' \in \Theta$, and $\omega' \in \Omega$ such that $x(\theta', \boldsymbol{\pi}_{i'}, \omega') > M'$ or $x(\theta', \boldsymbol{\pi}_{i'}, \omega') < -M'$.

First, we will consider the case where $x(\boldsymbol{\pi}_{i'}, \theta', \omega') < -M'$. Note that the expected revenue generated for any type θ must be bounded from below by $-v(|\Theta|)/\pi_{\min}$ if the mechanism guarantees non-zero expected revenue. This is because the maximum amount of expected revenue for any type can be at most $v(|\Theta|)$ or individual rationality will not be satisfied, and if expected revenue for any type is less than $-v(|\Theta|)/\pi_{\min}$, it is not possible to make up the revenue from other types. Further, this implies that in order for the mechanism to generate non-negative revenue, the bidder's expected utility for any type must be less than $v(|\Theta|) + v(|\Theta|)/\pi_{\min}$. Therefore, set

$$M' = \frac{v(|\Theta|)}{\pi_{\min}} + \frac{(1 - \epsilon_{\min})(2v(|\Theta|) + \frac{v(|\Theta|)}{\pi_{\min}} + 1)}{\epsilon_{\min}}$$

Then, the magnitude of the gradient of the hyper-plane defined by the affine combination of $x(\theta', \boldsymbol{\pi}_{i'}, \omega)$ for all $\omega \in \Omega$ must be at least:

$$||\nabla \boldsymbol{x}(\theta', \boldsymbol{\pi}_{i'}, \cdot)|| \geq \frac{(-\frac{v(|\Theta|)}{\pi_{\min}} + M')}{(1 - \epsilon_{\min})} = \frac{2v(|\Theta|) + \frac{v(|\Theta|)}{\pi_{\min}} + 1}{\epsilon_{\min}}$$

Let $\boldsymbol{z} \in \mathbb{R}^{|\Omega|}$ with $\sum_\omega z(\omega) = 0$ be the direction of the gradient of the hyper-plane defined by the lottery in the plane of the $|\Omega|$-simplex. Then by Lemma 5, there exists a $\boldsymbol{\pi}_j(\cdot|\theta_j)$ such that $(\boldsymbol{\pi}_{i'}(\cdot|\theta') - \boldsymbol{\pi}_j(\cdot|\theta_j)) \bullet \boldsymbol{z} > \epsilon_{\min}$. Then:

$$\begin{aligned}
&\sum_\omega \pi_j(\omega|\theta_j) U(\boldsymbol{\pi}_j, \theta_j, \boldsymbol{\pi}_j, \theta_j, \omega) \\
&\quad \geq \sum_\omega \pi_j(\omega|\theta_j) U(\boldsymbol{\pi}_j, \theta_j, \boldsymbol{\pi}_{i'}, \theta', \omega) \qquad \text{(by IC)} \\
&\quad \geq \sum_\omega \pi_j(\omega|\theta_j) U(\boldsymbol{\pi}_j, \theta_j, \boldsymbol{\pi}_{i'}, \theta', \omega) \\
&\qquad\quad - \sum_\omega \pi_{i'}(\omega|\theta') U(\boldsymbol{\pi}_{i'}, \theta', \boldsymbol{\pi}_{i'}, \theta', \omega) \quad \text{(by IR)} \\
&\quad \geq \sum_\omega (\pi_{i'}(\omega|\theta') - \pi_j(\omega|\theta_j)) x(\theta', \boldsymbol{\pi}_{i'}, \omega) - v(|\Theta|) \\
&\quad = (\boldsymbol{\pi}_{i'}(\cdot|\theta') - \boldsymbol{\pi}_j(\cdot|\theta_j)) \cdot \boldsymbol{z} ||\nabla \boldsymbol{x}(\theta', \boldsymbol{\pi}_{i'}, \cdot)|| - v(|\Theta|) \\
&\quad \geq \epsilon_{\min} ||\nabla \boldsymbol{x}(\theta', \boldsymbol{\pi}_{i'}, \cdot)|| - v(|\Theta|) \\
&\quad \geq v(|\Theta|) + v(|\Theta|)/\pi_{\min} + \epsilon_{\min}
\end{aligned}$$

Therefore, the seller cannot earn non-negative expected revenue for type $(\boldsymbol{\pi}_j, \theta_j)$, a contradiction.

It is straightforward to show that the combination of individual rationality and payments being bounded from below by $-\max\{M_{i^*<T}, M'\}$ implies that all payments must be bounded from above. We omit the details due to space considerations. Denote this upper bound by M''.

Therefore, let $M = \max\{M_{i^*<T}, M', M''\}$, and all payments are bounded by M. $\square$

With payments bounded, the final necessary result is the following stating that for any linear program where the variables for the set of optimal solutions is bounded, the corresponding sequence of linear programs is upper semi-continuous.

LEMMA 7 (MARTIN 1975 [12]). *Let $\boldsymbol{a}(\boldsymbol{t}), \boldsymbol{b}(\boldsymbol{t}), \boldsymbol{c}(\boldsymbol{t})$, and $\boldsymbol{d}(\boldsymbol{t})$ be vectors parameterized by the parameter vector $\boldsymbol{t} \in \mathcal{Q}$. Assume that $\boldsymbol{a}(\boldsymbol{t}), \boldsymbol{b}(\boldsymbol{t}), \boldsymbol{c}(\boldsymbol{t})$, and $\boldsymbol{d}(\boldsymbol{t})$ converge continuously to $\boldsymbol{a}(\boldsymbol{0}), \boldsymbol{b}(\boldsymbol{0}), \boldsymbol{c}(\boldsymbol{0})$, and $\boldsymbol{d}(\boldsymbol{0})$ as $\boldsymbol{t} \to \boldsymbol{0}$. Similarly, $\boldsymbol{A}(\boldsymbol{t})$, $\boldsymbol{B}(\boldsymbol{t})$, $\boldsymbol{C}(\boldsymbol{t})$, and $\boldsymbol{D}(\boldsymbol{t})$ are matrices that converge continuously to $\boldsymbol{A}(\boldsymbol{0})$, $\boldsymbol{B}(\boldsymbol{0})$, $\boldsymbol{C}(\boldsymbol{0})$, and $\boldsymbol{D}(\boldsymbol{0})$.*

Define the parameterized linear program $LP(\boldsymbol{t})$ as:

$$\begin{aligned}&\max_{\boldsymbol{x},\boldsymbol{q}} \quad \boldsymbol{c}'(\boldsymbol{t})\boldsymbol{x} + \boldsymbol{d}'(\boldsymbol{t})\boldsymbol{q}\\ &\textit{subject to}\\ &\boldsymbol{A}(\boldsymbol{t})\boldsymbol{x} + \boldsymbol{B}(\boldsymbol{t})\boldsymbol{q} = \boldsymbol{a}(\boldsymbol{t})\\ &\boldsymbol{C}(\boldsymbol{t})\boldsymbol{x} + \boldsymbol{D}(\boldsymbol{t})\boldsymbol{q} \leq \boldsymbol{b}(\boldsymbol{t})\\ &\boldsymbol{q} \geq \boldsymbol{0}\end{aligned}$$

If the set of optimal solutions of $LP(\boldsymbol{0})$, $\{(\boldsymbol{x},\boldsymbol{q}) : (\boldsymbol{x},\boldsymbol{q}) \in \arg\max(LP(\boldsymbol{0}))\}$, is bounded, then the objective value of $LP(\boldsymbol{t})$ is upper semi-continuous at $\boldsymbol{t} = \boldsymbol{0}$.

All of the pieces are now in place to prove our main result, that no IC and IR mechanism over the sequence can significantly outperform the optimal mechanism for the convergence point everywhere on the sequence.

Proof of Theorem 3. Note that the maximum revenue achievable for any given $\boldsymbol{\pi}_{i'}$ can be bounded from above by the following linear program:

$$\begin{aligned}&\max_{\boldsymbol{p},\boldsymbol{x}} \sum_{\theta}\sum_{\omega} \pi_{i'}(\theta,\omega)x(\theta,\boldsymbol{\pi}_{i'},\omega)\\ &\textit{subject to}\\ &\sum_{\omega} \pi_{i'}(\omega|\theta)U(\theta,\boldsymbol{\pi}_{i'},\theta,\boldsymbol{\pi}_{i'},\omega) \geq 0 \qquad \forall \quad \theta \in \Theta\\ &\sum_{\omega} \pi_{i'}(\omega|\theta)U(\theta,\boldsymbol{\pi}_{i'},\theta,\boldsymbol{\pi}_{i'},\omega)\\ &\qquad \geq \sum_{\omega} \pi_{i'}(\omega|\theta)U(\theta,\boldsymbol{\pi}_{i'},\theta',\boldsymbol{\pi}_{i'},\omega) \quad \forall \quad \theta,\theta' \in \Theta\\ &0 \leq p(\theta,\boldsymbol{\pi}_{i'},\omega) \leq 1 \qquad \forall \quad \theta \in \Theta, \omega \in \Omega\\ &-M \leq x(\theta,\boldsymbol{\pi}_{i'},\omega) \leq M \qquad \forall \quad \theta \in \Theta, \omega \in \Omega\end{aligned}$$

where the last constraint is a consequence of Lemma 6. Therefore, by Lemma 7, the objective of this program is upper semi-continuous at $\boldsymbol{\pi}^*$, and the result follows immediately. $\square$

Corollary 4 directly follows. The key insight is that any finite number of samples from the underlying distribution can be viewed as one signal from a more complicated distribution, and that this distribution still converges to a convergence point that will be IPV if the original convergence point is IPV.

Proof of Corollary 4. Let $\{(\theta_j,\omega_j)\}_{j=1}^N$ be a finite number of independent samples from the true distribution $\boldsymbol{\pi}_i$. Note the true distribution can be written as $\boldsymbol{\pi}_i = \boldsymbol{\pi}^* + \boldsymbol{\epsilon}_{\theta,i}$ for some $\boldsymbol{\epsilon}_{\theta,i} \in \mathbb{R}^{|\Omega|}$. Therefore, the probability of seeing samples $\{(\theta_j,\omega_j)\}_{j=1}^N$ and external signal ω is:

$$\begin{aligned}\boldsymbol{\pi}_i(\{(\theta_j,\omega_j)\}_{j=1}^N,\omega|\theta) &= \pi_i(\omega|\theta)\prod_{j=1}^N \pi_i(\omega_j|\theta_j)\pi(\theta_j)\\ &= (\pi^*(\omega|\theta) + \epsilon_{\theta,i}(\omega))\prod_{j=1}^N (\pi^*(\omega_j|\theta_j) + \epsilon_{\theta_j,i}(\omega_j))\pi(\theta_j)\end{aligned}$$

which converges to $\boldsymbol{\pi}^*(\{(\theta_j,\omega_j)\}_{j=1}^N,\omega|\theta)$ as $\boldsymbol{\pi}_i$ converges to $\boldsymbol{\pi}^*$. Moreover, the samples $\{(\theta_j,\omega_j)\}_{j=1}^N$ are independent of the final round's bidder type, so the optimal mechanism over the distribution $\boldsymbol{\pi}^*(\{(\theta_j,\omega_j)\}_{j=1}^N,\omega|\theta)$ is revenue equivalent to the optimal mechanism over $\boldsymbol{\pi}^*$. Therefore, a finite number of samples is equivalent to a higher dimensional signal, and Theorem 3 applies directly. $\square$

4. UNBOUNDING THE APPROXIMATION RATIO ON A CONVERGING SEQUENCE OF DISTRIBUTIONS

While Corollary 4 states that we can't learn a mechanism that guarantees optimal revenue, it leaves open the possibility that we can learn nearly optimal revenue. However, the following example demonstrates that the approximation ratio is unbounded in the number of bidder types.

Example 2. *Let the marginal distribution over the type of the bidder be given by $\pi(\theta) = 1/2^\theta$ for $\theta \in \{1,...,|\Theta|-1\}$ and $\pi(|\Theta|) = 1/2^{|\Theta|-1}$. Further let the value of the bidder for the item be $v(\theta) = 2^\theta$. Therefore, the expected value of the bidder's valuation is*

$$\sum_{\theta=1}^{|\Theta|-1}\left(\frac{1}{2^\theta}\right)2^\theta + \left(\frac{1}{2}\right)^{|\Theta|-1}2^{|\Theta|} = |\Theta| + 1$$

Assume that the external signal is binary, i.e. $\Omega = \{\omega_L,\omega_H\}$. Note that for a reserve price mechanism with a reserve price of $2^{|\Theta|}$, the expected revenue is 2. Further, if the distribution is IPV, this is the optimal mechanism [17].

As the following lemma shows, Example 2 gives a sequence of distributions that all satisfy the ACL condition and whose full surplus revenue grows without bound in the number of bidder types. However, the sequence converges to an IPV distribution that has constant revenue in the number of bidder types.

Lemma 8. *In the setting of Example 2, there exists a sequence of distributions $\{\boldsymbol{\pi}_i\}_{i=1}^\infty$ that converges to an IPV distribution and satisfies Assumption 1 such that for each distribution $\boldsymbol{\pi}_i$, their exists a mechanism $(\boldsymbol{p}_i,\boldsymbol{x}_i)$ whose expected revenue is $|\Theta|+1$.*

Proof. Let $\pi_i(\omega_L|\theta) = 1/2 + (1/i)(1/2 - (1/2)^{|\Theta|-\theta})$. Then, define the linear function

$$G_i(\boldsymbol{\pi}(\cdot|\theta)) = -\pi(\omega_L|\theta)i2^{|\Theta|} + i2^{|\Theta|-1} + 2^{|\Theta|-1}$$

Therefore, $G_i(\boldsymbol{\pi}_i(\cdot|\theta)) = 2^\theta$, the ACL condition, and by Theorem 2, for each $\boldsymbol{\pi}_i$, there exists a mechanism such that the expected revenue is $|\Theta|+1$. Furthermore, $\boldsymbol{\pi}_i(\cdot|\theta)$ converges to $\pi^*(\omega_L|\Theta) = 1/2$, an IPV distribution. Finally, $\pi_i(\omega_L|1) > 1/2$ while $\pi_i(\omega_L||\Theta|) < 1/2$, so the sequence satisfies Assumption 1. $\square$

Corollary 9. *The expected revenue generated by an IR and IC mechanism over an infinite sequence of distributions guarantees at best a $(|\Theta|+1)/(2+\epsilon)$ approximation to the revenue achievable by the optimal Bayesian IC and ex-interim IR mechanism if the distribution over types is exactly known. This is still true if the mechanism designer has access to a finite number of samples from the true distribution.*

5. APPROXIMATING THE OPTIMAL MECHANISM

Theorem 3 and Corollary 4 shows that we are unlikely to do well for distributions near IPV, but what if we can bound the true distribution away from IPV? Can we take advantage of correlation if there is enough correlation?

We now show that if the true distribution satisfies the ACL condition, then there is a single mechanism such that *for all distributions sufficiently close to the true distribution, the mechanism designer can do nearly as well as if she knew the true distribution.*

THEOREM 10. *For any distribution $\boldsymbol{\pi}^*$ that satisfies the ACL condition with optimal revenue R and given any positive constant $k > 0$, there exists $\epsilon > 0$ and a mechanism such that for all distributions, $\boldsymbol{\pi}'$, for which for all $\theta \in \Theta$, $||\boldsymbol{\pi}^*(\cdot|\theta) - \boldsymbol{\pi}'(\cdot|\theta)|| < \epsilon$, the revenue generated by the mechanism is greater than or equal to $R - k$.*

PROOF. By the assumption that there exists a mechanism that extracts full surplus for the distribution $\boldsymbol{\pi}^*$, there must be a mechanism that always allocates the item and leaves the bidder with an expected utility of 0. Let this mechanism be denoted by $(\boldsymbol{p}^*, \boldsymbol{x}^*)$. Note that this mechanism does not depend on a reported distribution, due to it being a mechanism over a single distribution. Let C be the value for the largest slope of the gradient of any lottery in the mechanism. Choose $\epsilon = k/(2C)$. Then the expected utility for any distribution $\boldsymbol{\pi}'(\cdot|\theta)$ with $||\boldsymbol{\pi}^*(\cdot|\theta) - \boldsymbol{\pi}'(\cdot|\theta)|| < \epsilon$ when optimally reporting $\theta' \in \Theta$ is bounded by:

$$-C\epsilon \leq \sum_{\omega} \pi'(\omega|\theta) U(\theta, \theta', \omega) \leq C\epsilon$$

Construct a new mechanism (not necessarily truthful) where all payments $x'(\theta, \omega) = x^*(\theta, \omega) - C\epsilon$ and set $p'(\theta, \omega) = p^*(\theta, \omega) = 1$. Then, the utility of the bidder for optimally misreporting is:

$$0 \leq \sum_{\omega} \pi'(\omega|\theta) U(\theta, \theta', \omega) \leq 2C\epsilon$$

which implies that the bidder always participates. Since the item is always allocated, the loss in revenue is equivalent to the gain in utility for the bidder. Therefore, the mechanism $(\boldsymbol{p}', \boldsymbol{x}')$ always guarantees revenue within $2C\epsilon = k$ of the optimal mechanism for any θ, so the expected revenue of the mechanism is greater than or equal to $R - k$. $\square$

Note that the proof of Theorem 10 is constructive. Theorem 10 is intuitively very reasonable, and likely what one would expect a-priori. For a class of distributions that are sufficiently close, there should be a mechanism that does about as well on all of them. However, we believe this result is a fundamental insight into the problem of correlated mechanism design; *the closeness necessary to achieve nearly full revenue is dependent on the magnitude of the largest gradient over all of the lotteries.*

Examining Figure 1c, it is intuitive that the greater the correlation, the smaller the gradient needs to be, i.e. if two points in Figure 1c are nearly on top of each other, the gradients of the lotteries necessary to distinguish between the two types is very large. Furthermore, if two valuations are far apart in magnitude (i.e. $v(\theta_1) \ll v(\theta_2)$), then the gradient of the lottery will be large. This intuition is proven formally in [2]. While formal conditions characterizing the relationship between correlation and the maximal revenue achievable for a given estimate is outside of the scope of this work, we believe that these results are suggestive of a path towards these conditions.

6. CONCLUSION

In this paper, we have presented the extremes of learning mechanisms for settings with correlated bidder distributions. On one hand, Theorem 3 and Corollary 4 suggest that learning is doomed in the worst case. On the other, Theorem 10 suggests that if we can get close to the true distribution we can do practically as well as if we knew the distribution. We interpret Theorem 3 and Corollary 4 as suggesting that getting close to the optimal revenue will be, effectively, impossible for distributions that are nearly IPV, while Theorem 10 implies that for a sufficiently correlated distribution we can likely do significantly better than IPV. The opportunity for mechanisms over distributions in between these two extremes is unknown and, we believe, a very interesting direction for future research.

Recent work [3] has employed automated mechanism design techniques in conjunction with methodologies from robust optimization to examine the practical performance of a particular technique for incorporating uncertainty into prior-dependent mechanism design. While this work does not focus on theoretical performance guarantees and the results are limited to a specific simulation setting, it is interesting to note that the degree of correlation is strongly related to the mechanisms ability to extract nearly optimal revenue, consistent with the intuition suggested by Theorem 10. Their work is able to achieve nearly optimal revenue by only focusing on portions of the set of possible distributions that are "reasonably likely" based on observed samples from the distribution and, effectively, ignoring the rest of the set of distributions. This again is consistent with the intuition that the mechanism designer must accept that the mechanism may not perform well on some subset of the distributions, Theorem 3, and that the mechanism designer must, therefore, optimize the mechanism for the likely distributions.

It seems likely that the way towards practical, implementable, provably efficient mechanisms with correlated bidder distributions will rely critically on the degree of correlation. However, the literature has traditionally ignored the degree of correlation, since full surplus extraction is possible with very little correlation [5, 7]. Our work suggests that the path forward will likely require three steps: First, the gradient of the lotteries implemented by the optimal mechanism must be bounded in order to determine how close the estimate must be to the true distribution. Second, the sample complexity of estimation procedures needed to ensure that the true distribution is close to the estimate must be analyzed. Third, mechanisms must be designed that do well on all distributions close to the estimate.

7. ACKNOWLEDGEMENTS

Albert and Conitzer are thankful for support from NSF under awards IIS-1527434 and CCF-1337215, ARO under grants W911NF-12-1-0550 and W911NF-11-1-0332, and a Guggenheim Fellowship.

This work has partially taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (CNS-1330072, CNS-1305287, IIS-1637736, IIS-1651089), ONR (21C184-01), and AFOSR (FA9550-14-1-0087). Stone serves on the Board of Directors of, Cogitai, Inc. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.

REFERENCES

[1] M. Albert, V. Conitzer, and G. Lopomo. Assessing the robustness of Cremer-McLean with automated mechanism design. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI '15, pages 763–769, Austin, TX, USA, January 2015. AAAI Press.

[2] M. Albert, V. Conitzer, and G. Lopomo. Maximizing revenue with limited correlation: The cost of ex-post incentive compatibility. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*, 2016.

[3] M. Albert, V. Conitzer, and P. Stone. Automated design of robust mechanisms. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, 2017.

[4] A. Blum, Y. Mansour, and J. Morgenstern. Learning valuation distributions from partial observation. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI '15, pages 798–805, Austin, TX, USA, 2015. AAAI Press.

[5] J. Cremer and R. P. McLean. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53(2):pp. 345–361, 1985.

[6] J. Cremer and R. P. McLean. Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):pp. 1247–1257, 1988.

[7] H. Fu, N. Haghpanah, J. Hartline, and R. Kleinberg. Optimal auctions for correlated buyers with sampling. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC '14, pages 23–36, New York, NY, USA, 2014. ACM.

[8] R. Gibbons. *Game Theory for Applied Economists.* Princeton University Press, 1992.

[9] Y. Kanoria and H. Nazerzadeh. Dynamic reserve prices for repeated auctions: Learning from bids. SSRN Paper. Suggested citation: Kanoria, Yash and Nazerzadeh, Hamid, Dynamic Reserve Prices for Repeated Auctions: Learning from Bids (July 1, 2014). Available at SSRN: http://ssrn.com/abstract=2444495 or http://dx.doi.org/10.2139/ssrn.2444495, July 2014.

[10] G. Kosenok and S. Severinov. Individually rational, budget-balanced mechanisms and allocation of surplus. *Journal of Economic Theory*, 140(1):126 – 161, 2008.

[11] G. Lopomo, L. Rigotti, and C. Shannon. Uncertainty in mechanism design. Oct. 2009.

[12] D. H. Martin. On the continuity of the maximum in parametric linear programming. *Journal of Optimization Theory and Applications*, 17(3):205–210, 1975.

[13] R. P. McAfee and P. J. Reny. Correlated information and mechanism design. *Econometrica*, 60(2):pp. 395–421, 1992.

[14] M. Mohri and A. M. Medina. Optimal regret minimization in posted-price auctions with strategic buyers. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 1871–1879. Curran Associates, Inc., 2014.

[15] M. Mohri and A. M. Medina. Learning algorithms for second-price auctions with reserve. *Journal of Machine Learning Research*, 3, 2015.

[16] J. Morgenstern and T. Roughgarden. The pseudo-dimension of nearly-optimal auctions. In *NIPS*, page Forthcoming, 12 2015.

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

[18] D. Pardoe, P. Stone, M. Saar-Tsechansky, T. Keskin, and K. Tomak. Adaptive auction mechanism design and the incorporation of prior knowledge. *Informs Journal on Computing*, 22(3):353–370, 2010.