Learning Optimal Auctions with Correlated Valuations from Samples

Chunxue Yang [1] **Xiaohui Bei** [1]

Abstract

In single-item auction design, it is well known due to Crémer and McLean that when bidders' valuations are drawn from a correlated prior distribution, the auctioneer can extract full social surplus as revenue. However, in most real-world applications, the prior is usually unknown and can only be learned from historical data. In this work, we investigate the robustness of the optimal auction with correlated valuations via sample complexity analysis. We prove upper and lower bounds on the number of samples from the unknown prior required to learn a $(1-\epsilon)$-approximately optimal auction. Our results reinforce the common belief that optimal correlated auctions are sensitive to the distribution parameters and hard to learn unless the prior distribution is well-behaved.

1. Introduction

As a means to facilitate efficient resource allocation, auctions are a fundamental tool in the modern economy and play a pivotal role in mechanism design theory. The classical auction design problems in economics usually assume a Bayesian setting (Myerson, 1981; Crémer & McLean, 1985), in which participants' valuations (types) are drawn from some prior probability distribution that is also common knowledge known to the auction designer. This prior distribution plays a vital role in the analysis of optimal auctions.

In the case of single-item auction design, Myerson (1981) gives precise characterization of the revenue-maximizing auction when bidders' private types are drawn from independent prior distributions. The Myerson auction assumes that the prior distributions of all bidders are known to the auctioneer. However, in most applications, the prior distributions are unknown and can only be learned from past data. This raises an intriguing question of how much data is sufficient and necessary to guarantee near-optimal expected revenue. Consequently, one active line of research applies a principled approach of designing auctions as a function of several samples drawn from an unknown distribution, with the goal of obtaining high revenue with future reports that are drawn from the same distribution. The number of samples required to approximate the optimal auction with regard to an unknown distribution is known as the *sample complexity* of the problem. The sample complexity of the Myerson auction was first studied by Cole & Roughgarden (2014). With a long line of follow-up works (Gonczarowski & Nisan, 2017; Syrgkanis, 2017; Huang et al., 2018), it was recently settled by Guo et al. (2019) with matching upper and lower bounds up to a poly-logarithmic factor.

There is another major limitation of the Myerson auction. The assumption that bidders' private values are drawn from *independent* prior distributions, as Myerson himself points out, is strong and does not hold in many real-world applications. For the general case of joint and possibly correlated prior distributions, a much stronger result than Myerson's is possible. More specifically, Crémer & McLean (1985; 1988) show that the auctioneer can extract full social surplus as revenue if the joint prior distribution satisfies a mild condition. However, this Crémer-McLean result is considered to be non-applicable in many practical settings, mainly due to its strong dependence on the common prior assumption. Therefore, it is important to understand how optimal auction design is sensitive to the errors in the Bayesian beliefs and to provide a robustness analysis for the single-item auction design problem with correlated prior distributions.

1.1. Our Results

In this work, we investigate the sample complexity of Crémer-McLean optimal auctions for bidders with correlated distributions. Our goal is to understand how many samples from a prior distribution are sufficient and necessary to learn an auction with revenue at least $(1-\epsilon)$ of the total social surplus. We provide upper and lower bounds for this sample complexity problem, which are summarized in Table 1.

Different from the Myerson auction, whose sample complexity is only a function of the number of bidders n and ϵ under the regularity assumption, we find that the sample

[1]School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. Correspondence to: Chunxue Yang <chunxue001@e.ntu.edu.sg>, Xiaohui Bei <xhbei@ntu.edu.sg>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

Table 1. The sample complexity bounds of single-bidder Crémer-McLean auction. (K, α, and η represent the size of the valuation type set, the degree of correlation among bidders, and the smallest marginal probability of any valuation type, respectively.)

UPPER BOUND	$\tilde{O}(K^2\eta^{-3}\alpha^{-2}\epsilon^{-2})$
LOWER BOUND	$\Omega(K\eta^{-1}\alpha^{-2}\epsilon^{-2})$

complexity of Crémer-McLean also sensitively depends on several other key characteristics of the joint valuation type distribution. These factors include: (1) the size of valuation type set; (2) the degree of correlation among the bidders; and (3) the smallest marginal probability of any valuation type. As a direct corollary, our result suggests that when the level of correlation among the bidders is low, or there is some valuation with very small marginal probability, learning a nearly optimal auction becomes impossible without a lot of samples.

We prove the sample complexity upper bound by presenting a computationally efficient learning algorithm that achieves this sample complexity. The algorithm follows a rather intuitive idea. We first construct from the samples an empirical prior distribution. Then we shift down each valuation slightly. This step is to ensure interim individual rationality without losing too much revenue. Finally we simply apply the Crémer-McLean auction with regard to the empirical distribution and the down-shifted valuations.

For the lower bound, we follow the framework of Guo et al. (2019) and construct a family of distributions, such that any algorithm that achieves $(1-\epsilon)$ of the optimal revenue must be able to distinguish between them. We construct these distributions to be very close to each other in terms of Kullback-Leibler divergence. It therefore requires many samples to distinguish between them.

1.2. Related Work

Besides Myerson auction, sample complexity has also been used to analyze the robustness of many other types of auctions, such as t-level auctions (Morgenstern & Roughgarden, 2015), auctions with side information (Devanur et al., 2016), second-price auctions with anonymous reserves (Jin et al., 2019), multi-item auctions (Brustle et al., 2020), and non-truthful auctions (Hartline & Taggart, 2019).

Sample complexity has been considered in the context of correlated auctions before. Fu et al. (2014) showed that if there is a finite set of distributions from which the true distribution can be drawn, then the sample complexity for the Crémer-McLean auction is of the same order as the number of possible distributions. Note that their result only applies to the case of a finite set of distributions, while we do not have such restrictions. In another work, Albert et al. (2017a) studied the sample complexity of a Bayesian mechanism design paradigm that guarantees an additive approximation to full surplus revenue. The difference of our work from theirs is that we follow the setting of Crémer & McLean (1985) and require auctions to satisfy the constraint of ex-post incentive compatibility instead of the Bayesian one. In addition, we provide a lower bound of the sample complexity in this paper, which is important for understanding full surplus extraction.

The robustness of Crémer-McLean auction has also been studied from other directions. The genericity of full surplus extraction is discussed in (Heifetz & Neeman, 2006; Barelli, 2009; Chen & Xiong, 2013). Moreover, many works have assessed how sensitive the Crémer-McLean auction is to the relaxation of the technical assumptions, such as risk neutrality or unlimited liability (Robert, 1991), absence of cooperation among buyers (Laffont & Martimort, 2000), lack of competition among sellers (Peters, 2001), and uniqueness of each valuation's conditional distribution over the signals (Albert et al., 2015).

2. Preliminaries

To understand Crémer-McLean auction, we usually consider the case of a single bidder with an external signal, which captures the most important aspects of an auction with correlated bidders. We show in Section 5 how to generalize the single-bidder results to the case with multiple bidders.

We assume the bidder has a valuation type θ drawn from a finite set of discrete types $\Theta = \{1, 2, \ldots, K\}$ with $K = |\Theta|$ denoting the number of possible types. This valuation type θ is the bidder's private information. In addition, the bidder has a valuation function $v : \Theta \to \mathbb{R}_+$ that maps types to valuations of the indivisible item for sale. The types' valuations are assumed to be increasing, i.e., $v(1) < v(2) < \cdots < v(K)$. There is an external signal ω drawn from some discrete signal set Ω, which is correlated with the bidder's valuation type. We assume $|\Omega| \geq K$ in this paper. The external signal ω is known to the auctioneer. The joint distribution over the valuation types and external signals is denoted as $\pi(\omega, \theta)$, and the marginal distribution over the bidder's type θ is denoted as $\pi(\theta)$. This single bidder and external signal setting has also been adopted in several other works on correlated auctions (McAfee & Reny, 1992; Albert et al., 2015; 2016; 2017a;b).

Auctions. In an auction, the bidder with private valuation type θ reports a (possibly different) type $\theta' \in \Theta$, and the auctioneer receives θ' and an external signal ω. An auction $M(x, p)$ consists of two functions x and p, both taking θ' and ω as input. The allocation function $x : \Theta \times \Omega \to \{0, 1\}$ decides whether the bidder wins the item. The payment function $p : \Theta \times \Omega \to \mathbb{R}$ specifies the amount of money that the bidder needs to pay to the auctioneer. We assume quasi-linear utility, that is, the bidder's utility is given by

$$u(\theta, \theta', \omega) = v(\theta)x(\theta', \omega) - p(\theta', \omega).$$

In this work, we are only interested in auctions that satisfy the following two desirable properties.

Definition 1 (Ex-post Incentive Compatibility)**.** *An auction is ex-post incentive compatible (IC) if for any realization of the valuation type θ and external signal ω, the bidder's utility is maximized when she reports her true type. That is,*

$$v(\theta)x(\theta, \omega) - p(\theta, \omega) \geq v(\theta)x(\theta', \omega) - p(\theta', \omega) \quad \forall \theta' \in \Theta.$$

We focus on ex-post IC auctions following the setting of Crémer & McLean (1985) and subsequent works (McAfee & Reny, 1992; Fu et al., 2014). Note that this property is relaxed to Bayesian IC in (Albert et al., 2016; 2017a;b). On the one hand, ex-post IC ensures the bidder never regrets reporting her type truthfully, no matter what the realization of the external signal is. Thus it makes the auction more robust than that with Bayesian IC. On the other hand, Bayesian IC auctions can extract full surplus as revenue for a larger set of distributions.

Definition 2 (Interim Individual Rationality)**.** *An auction is interim individual rational (IR) if for any realization of the valuation type θ, the bidder's expected utility is always non-negative when she reports the truth. That is,*

$$\mathbb{E}_{\omega \sim \pi(\theta)}[v(\theta)x(\theta, \omega) - p(\theta, \omega)] \geq 0.$$

We focus on revenue maximization as our objective. That is, the goal of the auction designer is to maximize the expected revenue the auction derives. We denote by $\textsc{Rev}(M, \pi)$ the expected revenue of auction M when ω, θ is drawn from joint distribution π. The auction that maximizes $\textsc{Rev}(\cdot, \pi)$ is called the *optimal auction*, and its revenue is denoted by $\textsc{Opt}(\pi)$.

Crémer-McLean optimal auction. Crémer & McLean (1985) characterizes the optimal auction for joint distributions that meet a mild condition:

Definition 3 (Crémer-McLean Condition)**.** *For a joint distribution π, let Γ be the following matrix whose rows are indexed by the elements of Θ, and columns are indexed by the elements of Ω:*

$$\Gamma = \begin{bmatrix} \pi(\omega_1|1) & \pi(\omega_2|1) & \cdots & \pi(\omega_{|\Omega|}|1) \\ \pi(\omega_1|2) & \pi(\omega_2|2) & \cdots & \pi(\omega_{|\Omega|}|2) \\ \vdots & & & \\ \pi(\omega_1|K) & \pi(\omega_2|K) & \cdots & \pi(\omega_{|\Omega|}|K) \end{bmatrix}. \quad (1)$$

The distribution π is said to satisfy Crémer-McLean condition if Γ has rank K.

Under Crémer-McLean condition, there always exists $\boldsymbol{p} = (p(\omega))_{\omega \in \Omega}$ such that for each $\theta \in \Theta$, $\sum_{\omega \in \Omega} \pi(\omega|\theta)p(\omega) = v(\theta)$. The Crémer-McLean optimal auction always allocates the item to the bidder, and charges her a payment of $p(\omega)$. This auction can extract the full social surplus $\sum_{\theta \in \Theta} \pi(\theta)v(\theta)$ as revenue, and satisfies ex-post IC and interim IR.

The Crémer-McLean condition is sufficient for the existence of an ex-post IC full surplus extraction auction. For Bayesian IC auctions, the condition for full surplus extraction can be further relaxed to what is known as the Albert-Conitzer-Lopomo condition (2016).

Sample complexity. Assume that the joint distribution π is unknown to us (the auctioneer), and we can access π only in the form of i.i.d. samples drawn from it. The *sample complexity* is defined as the asymptotically smallest number m such that given m i.i.d. samples from a joint distribution π, there exists an algorithm that can learn an auction M achieving revenue $\textsc{Rev}(M, \pi) \geq (1 - \epsilon)\textsc{Opt}(\pi)$ with high probability.

Degree of correlation. Our upper and lower bounds on the sample complexity both depend on the degree of correlation between the valuation types and external signals. To capture this notion quantitatively, we introduce the following parameterized condition.

Definition 4 (α-strongly correlated distribution)**.** *The distribution π is said to be α-strongly correlated if the singular values of any $K \times K$ nonsingular submatrix of Γ defined in Eq.1 are at least α.*

α-strong correlation interpolates between independence and full-correlation (i.e., determinism). On one end, if the external signal is independent of the bidder's type, then any $K \times K$ submatrix has singular value 0, and thus we consider the joint distribution π as 0-strongly correlated. On the other end, if in each row of Γ, one entry is 1 and the others are 0, which means θ can decide ω, then the singular values of any $K \times K$ nonsingular submatrix of Γ are 1, which means π is 1-strongly correlated. In the following sections, we assume the prior distribution is α-strongly correlated ($0 < \alpha \leq 1$).

Note that Albert et al. (2017a) also provides a parameterized condition, known as γ-separation, to measure the degree of correlation. However, α-strong correlation and γ-separation are distinct notions (see an example in the Supplementary Material for their distinction).

Regularity assumption. We take allowable distributions in this paper to be the ones that satisfy $\pi(\theta) \geq \eta$ for all $\theta \in \Theta$, where η is a constant in $(0, \frac{1}{K}]$.

This regularity assumption is necessary because otherwise, if the marginal probabilities are allowed to be arbitrarily close to zero, a nearly full surplus extraction is impossible. Here we give an informal example. For any algorithm that uses m $(m \geq 2)$ samples, consider a type θ with marginal probability $\pi(\theta) = \frac{1}{m}$. In addition, type θ has a very high valuation such that $\pi(\theta)v(\theta)$ accounts for a constant proportion in social surplus. Note that with probability at least $\frac{1}{4}$, no sample of type θ will appear in the m samples. Due to the lack of information on type θ, the algorithm cannot return a nearly full surplus extraction with high probability.

3. The Upper Bound

We present in this section an algorithm and its analysis that achieve the following sample complexity upper bound. The algorithm works for all α-strongly correlated distributions whose marginal probabilities are at least η.

Theorem 1. *For any $0 < \epsilon < 1$ and any α-strongly correlated distribution π whose marginal probabilities are at least η, Algorithm 2 returns an auction with expected revenue at least $(1-\epsilon)$ of the full surplus with probability at least $1-\delta$, if the number of samples m satisfies*

$$m \geq 90 \cdot K\eta^{-3}\alpha^{-2}\epsilon^{-2} \cdot \max\{5\ln(6K\delta^{-1}), 8K\}. \quad (2)$$

Algorithm 2 is not complicated. We first construct from the samples an empirical prior distribution $\hat{\pi}$. Then we decrease each valuation by a small value. This step is to ensure interim IR without losing too much revenue. Finally, we apply the Crémer-McLean auction with regard to the empirical distribution and the down-shifted valuations. We call Algorithm 2 the empirical Crémer-McLean auction[1].

The remaining of this section is dedicated to the proof of Theorem 1. Intuitively, we will show the following in order:

1. with enough samples, the empirical matrix $\hat{\Gamma}'$ defined in Eq. 3 will be very close to Γ', where Γ' represents the matrix whose entries are indexed by the same way as $\hat{\Gamma}'$ but constituted by the true probabilities;
2. the feasibility of finding a subset Ω' at Step 2 of Algorithm 1;
3. with high probability, Algorithm 2 returns an auction that is ex-post IC and interim IR and can generate nearly optimal revenue.

3.1. Algorithm Analysis

Our algorithm relies on constructing from the samples an empirical matrix $\hat{\Gamma}'$ and a shift-down valuation $\boldsymbol{v}'$, which are close to the true matrix Γ' and true valuation $\boldsymbol{v}$. To upper bound their differences, we first need the following two concentration inequalities.

Lemma 1 (e.g., see (Mitzenmacher & Upfal, 2017))**.** *Let X be a binomial (n,p) random variable. For all $t \in (0,1)$,*

[1] Formally speaking, this algorithm is a slight modification of the empirical Crémer-McLean auction because of the down-shifted valuations. In this work, we will still call it the empirical Crémer-McLean auction for notation convenience.

Algorithm 1 Learning the bidder's payment in an empirical Crémer-McLean auction

Input: the empirical joint distribution $\hat{\boldsymbol{\pi}}$ and the down-shifted valuations $v'(\theta)$
Output: the bidder's payment function $\hat{p}(\omega)$
Step 1. Obtain the empirical matrix $\hat{\Gamma}$ defined in Eq. 1 by letting $\hat{\pi}(\omega|\theta) = \hat{\pi}(\omega,\theta)/\sum_{\bar{\omega}\in\Omega}\hat{\pi}(\bar{\omega},\theta)$.
Step 2. Find a subset Ω' of Ω with size K, such that the minimum singular value of the $K \times K$ matrix $\hat{\Gamma}'$

$$\begin{bmatrix} \hat{\pi}(\omega_1'|1) & \hat{\pi}(\omega_2'|1) & \cdots & \hat{\pi}(\omega_K'|1) \\ \hat{\pi}(\omega_1'|2) & \hat{\pi}(\omega_2'|2) & \cdots & \hat{\pi}(\omega_K'|2) \\ \vdots & & & \\ \hat{\pi}(\omega_1'|K) & \hat{\pi}(\omega_2'|K) & \cdots & \hat{\pi}(\omega_K'|K) \end{bmatrix} \quad (3)$$

is greater than $2\alpha/3$.
Step 3. Solve the following system of equations for $\hat{p}(\omega)$ $(\omega \in \Omega')$

$$\sum_{\omega\in\Omega'} \hat{\pi}(\omega|\theta)\hat{p}(\omega) = v'(\theta) \quad \forall \theta \in \Theta \quad (4)$$

and let $\hat{p}(\omega) = 0$ for all ω not in Ω'.
Step 4. Return the bidder's payment function $\hat{p}(\omega)$.

Algorithm 2 Empirical Crémer-McLean Auction

Input: m i.i.d. samples from the distribution π
Output: an auction that decides the allocation and payment
Step 1. Let $\hat{\pi}$ be the empirical joint distribution, i.e., the uniform distribution over the samples.
Step 2. Let

$$v'(\theta) = \max\left\{0, v(\theta) - \frac{3\sqrt{K}\,\|\boldsymbol{v}\|\,\zeta}{2\alpha}\right\},$$

where $\|\boldsymbol{v}\|$ is the Euclidean norm of the vector $\boldsymbol{v} = (v(1), v(2), \ldots, v(K))$ and

$$\zeta = \sqrt{\frac{10}{\eta m}} \max\left\{\sqrt{5\ln(6K\delta^{-1})}, \sqrt{8K}\right\}.$$

Step 3. Run Algorithm 1 with the empirical distribution π and down-shifted valuations $v'(\theta)$ as input to get the payment function $\hat{p}(\omega)$.
Step 4. Return an auction that always allocates the item to the bidder and charges her $\hat{p}(\omega)$.

we have

$$\Pr[X < (1-t)np] \le \exp\left(-\frac{npt^2}{2}\right).$$

Lemma 2 (e.g., see (Devroye et al., 1983)). *Let $(X_1, \cdots, X_k)$ be a multinomial $(n, p_1, \cdots, p_k)$ random vector. For all $t \in (0,1)$ and all k satisfying $\frac{k}{n} \le \frac{t^2}{20}$, we have*

$$\Pr\left[\sum_{i=1}^{k} |X_i - \mathbb{E}(X_i)| > nt\right] \le 3\exp\left(-\frac{nt^2}{25}\right).$$

The above concentration inequalities allows us to formalize the intuition that $\hat{\Gamma}'$ is close to Γ' with the following lemma.

Lemma 3. *If the number of samples m satisfies Eq. 2, with probability at least $1-\delta$, for any subset Ω' of Ω with size K and for all $\theta \in \Theta$, we have*

$$\sum_{\omega \in \Omega'} |\hat{\pi}(\omega|\theta) - \pi(\omega|\theta)| \le \zeta. \tag{5}$$

The proof of Lemma 3, together with other omitted proofs, can be found in the Supplementary Material.

Next we analyze the auction returned by Algorithm 2 under the assumption that the conclusion of Lemma 3 holds.

Feasibility of Algorithm 1. We show that at Step 2 in Algorithm 1, we can find such a feasible subset Ω'. Equivalently, we only need to show that there exists a submatrix $\hat{\Gamma}'$ of the empirical matrix $\hat{\Gamma}$, whose singular values are greater than $\frac{2\alpha}{3}$. We will rely on Neumann series to prove this claim.

Lemma 4 (Neumann Series e.g., see (Schechter, 1996)). *Suppose A is an $n \times n$ matrix. If $\|A\| < 1$, then $I - A$ is invertible and its inverse is the series*

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k,$$

where I represents the identity matrix, and $\|\cdot\|$ stands for the operator norm.

Lemma 5. *At Step 2 of Algorithm 1, there exists a subset Ω' of Ω with size K, such that the minimum singular value of the submatrix $\hat{\Gamma}'$ defined in Eq. 3 is greater than $\frac{2}{3}\alpha$.*

IC and IR of the auction. The ex-post IC of the auction returned by Algorithm 2 is rather straightforward. In fact, no matter what type the bidder reports, she will always get the item, and the payment only depends on the external signal but not on this bidder's reported type. Thus ex-post IC always holds deterministically for the returned auction.

To show the returned auction is interim IR, it suffices to show that for any type $\theta \in \Theta$, the bidder's expected payment is no greater than her valuation.

Lemma 6. *In the auction returned by Algorithm 2, for all $\theta \in \Theta$, the expected payment of the bidder with type θ satisfies*

$$v'(\theta) - \frac{3\sqrt{K}\,\|\boldsymbol{v}\|\,\zeta}{2\alpha} \le \sum_{\omega \in \Omega'} \pi(\omega|\theta)\hat{p}(\omega) \le v(\theta).$$

Near-optimal revenue. Finally we show the near-optimal revenue guarantee of our auction. Lemma 6 also shows that the expected revenue loss $\Delta(\pi)$ of the auction returned by Algorithm 2 is upper bounded by $\frac{3\sqrt{K}\|\boldsymbol{v}\|\zeta}{\alpha}$. Therefore, when m satisfies Eq. 2, with probability at least $1-\delta$ we have

$$\frac{\Delta(\pi)}{\textsc{Opt}(\pi)} \le \frac{\Delta(\pi)}{\eta \cdot \sum_{\theta\in\Theta} v(\theta)} \le \frac{3\sqrt{K}\zeta}{\alpha\eta} \cdot \frac{\|\boldsymbol{v}\|}{\sum_{\theta\in\Theta} v(\theta)} \le \epsilon.$$

This completes the proof of Theorem 1.

Computational complexity. Algorithm 2 can be easily implemented in time polynomial in the number of bidder types, external signals and i.i.d. samples.

It is clear that in Algorithm 2, Step 1, 2 and 4 each takes polynomial (linear) time. For Algorithm 1, at Step 2 we can initialize Ω' to be empty, then add one new ω column at a time, and check whether the first $|\Omega|$ singular values of $\{\hat{\pi}(\omega|\cdot) : \omega \in \Omega'\}$ are greater than $\frac{2\alpha}{3}$ via singular value decomposition after each addition. We repeatedly add columns until the size of Ω' reaches K. These operations can be implemented in $O(K^3|\Omega|)$ time. At Step 3 of Algorithm 1, it also takes at most $O(K^3)$ time to solve the linear system (e.g., see (Barrodale & Stuart, 1981)) and compute $\hat{p}(\omega)$.

4. The Lower Bound

In this section, we discuss the sample complexity lower bound for any near-optimal auction. Our main result is a lower bound that matches the upper bound in previous section up to a factor of $\eta^{-2}\max\{\ln(K\delta^{-1}), K\}$. We assume $K \ge 3$, $\alpha < \frac{1}{\sqrt{K}}$, and $\epsilon < \frac{\eta K}{300}$ in the below theorem and its analysis.

Theorem 2. *Suppose an algorithm A, given m independent samples from an unknown α-strongly correlated distribution with marginal probabilities at least η, returns an auction with expected revenue at least $(1-\epsilon)$ of the social surplus with probability at least 0.99. Then m must be at least $\Omega(K\eta^{-1}\alpha^{-2}\epsilon^{-2})$.*

Remark. Albert et al. (2017b) proves that if the bidder's type and the external signal are approximately independent, no algorithm can return a Bayesian IC and IR auction that guarantees $(1-\epsilon)$ of the optimal revenue with any finite number of samples. The argument is proven for the more general Bayesian auctions, and therefore holds for the ex-post case as well. For ex-post IC auctions, this result can be directly implied by Theorem 2 when α is infinitely small.

Before further analysis, we first provide some discussion on the format of a near-optimal auction and on distribution classification, which will be useful for our proof of the lower bound theorem.

Format of any near-optimal auction. We would like to first clarify what nearly optimal auctions are like. The following lemma indicates that if the revenue of an auction is very close to the optimal one, then the auction must take the same form as Crémer-McLean auction.

Lemma 7. *For any given (π, v, Ω) and a small enough ϵ, suppose $M(x, p)$ is an ex-post IC and interim IR auction which can extract at least $(1-\epsilon)$ of the social surplus as revenue. Then (1) the allocation rule x must always allocate the item to the single bidder, and (2) the payment rule p is a function which only depends on the external signals.*

In the following analysis, without loss of generality, we only consider the auctions taking the form in Lemma 7.

Distribution classification. We say a classification algorithm $A : S^m \to \{\boldsymbol{\rho}, \boldsymbol{\varphi}\}$ distinguishes $\boldsymbol{\rho}$ and $\boldsymbol{\varphi}$ correctly with m samples, if for any $\pi \in \{\boldsymbol{\rho}, \boldsymbol{\varphi}\}$, $A(s_1, s_2, \ldots, s_m) = \pi$ with probability at least $\frac{2}{3}$, where $s_1, s_2, \ldots, s_m$ are i.i.d. samples from π. We use the following connection between Kullback-Leibler (KL) divergence of two distributions and the number of samples needed to distinguish them.

Lemma 8 (e.g., see (Huang et al., 2018)). *Suppose there is a classification algorithm that distinguishes $\boldsymbol{\rho}$ and $\boldsymbol{\varphi}$ correctly with m samples. Then, the number of samples m is at least:*

$$\Omega(D_{\mathsf{SKL}}(\boldsymbol{\rho}, \boldsymbol{\varphi})^{-1}),$$

where $D_{\mathsf{SKL}}(\boldsymbol{\rho}, \boldsymbol{\varphi})$ means the symmetric version of KL divergence between $\boldsymbol{\rho}$ and $\boldsymbol{\varphi}$, which is defined as the sum of $D_{\mathsf{KL}}(\boldsymbol{\rho}\|\boldsymbol{\varphi})$ and $D_{\mathsf{KL}}(\boldsymbol{\varphi}\|\boldsymbol{\rho})$.

4.1. Construction of Hard Instances

To prove the lower bound theorem, we will construct a class $\mathcal{H}$ of distributions. Our plan is to show that any algorithm that has a good enough revenue approximation on all distributions in $\mathcal{H}$ must take a lot of samples. Then what properties do we need from distributions in $\mathcal{H}$? Intuitively, on the one hand, we would like these distributions to be very similar, so that it would take many samples to distinguish between them. On the other hand, they are supposed to be different in the sense that they cannot share an auction that gives an approximately optimal expected revenue with high probability for all of them. This idea has also been used in several other works on sample complexity lower bounds (Cole & Roughgarden, 2014; Huang et al., 2018; Guo et al., 2019).

Our construction is presented as follows. Let

$$\mathcal{H} = \{\boldsymbol{\pi}_S : S \subseteq \{1, \ldots, K-2\}\},$$

where $\boldsymbol{\pi}_S$ is defined as

$$\begin{aligned}
&\text{for } i = 1, 2, \ldots, K-2, \\
&\boldsymbol{\pi_S}(\omega_i, i) = \begin{cases} \left(\dfrac{1+\sqrt{K}\alpha}{2}\right)\eta & i \notin S \\ \left(\dfrac{1+\sqrt{K}\alpha}{2} + \epsilon'\right)\eta & i \in S \end{cases} \\
&\boldsymbol{\pi_S}(\omega_K, i) = \begin{cases} \left(\dfrac{1-\sqrt{K}\alpha}{2}\right)\eta & i \notin S \\ \left(\dfrac{1-\sqrt{K}\alpha}{2} - \epsilon'\right)\eta & i \in S \end{cases} \\
&\boldsymbol{\pi_S}(\omega_{K-1}, K-1) = \frac{(1+\sqrt{K}\alpha)[1-(K-1)\eta]}{2}, \\
&\boldsymbol{\pi_S}(\omega_K, K-1) = \frac{(1-\sqrt{K}\alpha)[1-(K-1)\eta]}{2}, \\
&\boldsymbol{\pi_S}(\omega_{K-1}, K) = \frac{(1-\sqrt{K}\alpha)\eta}{2}, \\
&\boldsymbol{\pi_S}(\omega_K, K) = \frac{(1+\sqrt{K}\alpha)\eta}{2}.
\end{aligned}$$

in which we set $\epsilon' = \frac{125\alpha\epsilon}{\sqrt{K}}$.

In other words, $\mathcal{H}$ contains 2^{K-2} distributions. The matrix Γ for each distribution $\pi_S \in \mathcal{H}$ is like

$$\begin{bmatrix}
\frac{1+\sqrt{K}\alpha}{2}(+\epsilon') & 0 & \ldots & 0 & \frac{1-\sqrt{K}\alpha}{2}(-\epsilon') \\
0 & \frac{1+\sqrt{K}\alpha}{2}(+\epsilon') & \ldots & 0 & \frac{1-\sqrt{K}\alpha}{2}(-\epsilon') \\
\vdots & & & & \\
0 & 0 & \ldots & \frac{1+\sqrt{K}\alpha}{2} & \frac{1-\sqrt{K}\alpha}{2} \\
0 & 0 & \ldots & \frac{1-\sqrt{K}\alpha}{2} & \frac{1+\sqrt{K}\alpha}{2}
\end{bmatrix}.$$

The minimum singular value of the matrix above is at least α (see the Supplementary Material for reason). Hence by definition, the distributions in $\mathcal{H}$ are all α-strongly correlated. Besides, the marginal probabilities are at least η for all $\pi \in \mathcal{H}$.

Next we set the valuation of each type. In order to make the learning of optimal auction for this family of distributions as difficult as possible, we set the valuation of type K to be much larger than that of other types. Specifically, we set

$$\begin{aligned}
v(i) &= \frac{25\epsilon + o(\epsilon)}{\eta(K-2)} \text{ for } i = 1, 2, \ldots, K-2, \\
v(K-1) &= \frac{\epsilon}{1-(K-1)\eta} + v(K-2), \\
v(K) &= \frac{1 - \sum_{i\in[K-1]} \pi(i)v(i)}{\eta}.
\end{aligned} \tag{6}$$

Then for all $\pi \in \mathcal{H}$, the revenue of optimal auction is $\textsc{Opt}(\pi) = \sum_{\theta\in\Theta} \pi(\theta)v(\theta) = 1$.

4.2. Analysis of Hard Instances

Fix any $1 \le i \le K-2$ and any $S' \subseteq \{1,\dots,K-2\}\backslash\{i\}$, let $\pi^1 = \pi_{S'}$ and $\pi^2 = \pi_{S'\cup\{i\}}$ be a pair of distributions in $\mathcal{H}$ that are identical except for $\pi(\cdot|i)$. That is, we have

$$\pi^1(\omega_i, i) = \tfrac{(1+\sqrt{K}\alpha)\eta}{2} \qquad \pi^1(\omega_K, i) = \tfrac{(1-\sqrt{K}\alpha)\eta}{2}$$
$$\pi^2(\omega_i, i) = (\tfrac{1+\sqrt{K}\alpha}{2} + \epsilon')\eta$$
$$\pi^2(\omega_K, i) = (\tfrac{1-\sqrt{K}\alpha}{2} - \epsilon')\eta.$$

In this part, our aim is to prove that to construct an auction that generates an approximately optimal expected revenue w.r.t. an unknown distribution in $\mathcal{H}$, the algorithm must be able to distinguish between π^1 and π^2. The sample complexity of learning near-optimal auction thus boil down to that of distribution classification, for which we have rather mature tools and results.

We formalize our goal with the following lemma.

Lemma 9. *Suppose an algorithm takes m samples from an arbitrary distribution $\pi \in \mathcal{H}$ and returns, with probability at least 0.99, an auction whose expected revenue is at least $(1-\epsilon)\textsc{Opt}(\pi)$. Then the number of samples m is at least $\Omega(D_{\mathsf{SKL}}(\pi^1,\pi^2)^{-1})$.*

Proof sketch. To prove this lemma, we first suppose by contradiction that the algorithm, denoted as A, takes $m < c \cdot D_{\mathsf{SKL}}(\pi^1,\pi^2)^{-1}$ samples from an unknown distribution π, for some sufficiently small constant c, and output an auction $A(\pi)$. That means this algorithm cannot distinguish π^1 and π^2. We will then show that there exists a distribution in $\mathcal{H}$ such that the expected revenue loss due to the mistakes made on distribution classification is at least ϵ. In particular, we show the following in order.

- For each $1 \le i \le K-2$, when the bidder's type is i, there is a constant fraction of the distributions in $\mathcal{H}$ with which the expected revenue loss of auction $A(\pi)$ is $\Omega(\epsilon/(\eta K))$.
- By a counting argument, we can find some distribution π^* in $\mathcal{H}$, such that there are $\Omega(K)$ bidder's types with which auction $A(\pi^*)$ suffers from an expected revenue loss of $\Omega(\epsilon/(\eta K))$.
- The expected revenue loss of auction $A(\pi^*)$ with distribution π^* is then
$$\sum_{\theta\in\Theta} \pi^*(\theta)\cdot(\text{revenue loss when bidder's type is }\theta)$$
$$\ge \eta\cdot\Omega(K)\cdot\Omega\left(\frac{\epsilon}{\eta K}\right) = \Omega(\epsilon).$$
- Finally we convert the above claim on the expected revenue loss to a with-high-probability claim. That is, we prove that with constant probability (at least 0.01), the expected revenue loss of auction $A(\pi^*)$ will still be $\Omega(\epsilon)$.

$\square$

In the remaining of this section, we formalize this proof. Let
$$\Delta_{M,\pi}(\theta) = v(\theta) - \sum_{\omega\in\Omega} \pi(\omega|\theta)p(\omega)$$
denote the expected revenue loss of auction M conditioned on that the bidder's type is θ. We define $\mathcal{M}$ as the following set of auctions:

$$\mathcal{M} = \Big\{M:\ \text{For all } \pi\in\mathcal{H},\ \Delta_{M,\pi}(K) \le \frac{\epsilon}{\eta} \text{ and } \Delta_{M,\pi}(K-1) \le \frac{\epsilon}{1-(K-1)\eta}\Big\}.$$

The auction returned by A, denoted as $A(\pi)$, belongs to $\mathcal{M}$ with probability at least 0.99, as a $(1-\epsilon)$ revenue approximation must be in $\mathcal{M}$. Next, we define two subsets of $\mathcal{M}$: for $j \in \{1,2\}$,

$$\mathcal{M}^j = \Big\{M \in \mathcal{M} : \Delta_{M,\pi^j}(i) \le \frac{25\epsilon}{\eta K} \text{ and } M \text{ is interim IR for } \pi^j\Big\}.$$

Actually these two subsets of $\mathcal{M}$ have no intersection, which means that there is no auction that can extract almost full surplus for both π^1 and π^2 when the bidder has type i.

Lemma 10. *We have $\mathcal{M}^1 \cap \mathcal{M}^2 = \varnothing$.*

With Lemma 10, we can now formalize the intuition that A takes too few samples to make different decisions on π^1 and π^2, and therefore must have a large revenue loss on at least one of them.

Lemma 11. *For either $j=1$ or $j=2$ (or both), we have*

$$\Pr_{A(\pi^j)}\left[\Delta_{A(\pi^j),\pi^j}(i) > \frac{25\epsilon}{\eta K}\right] > \frac{3}{10}. \tag{7}$$

Next we argue that there must exist many bidder types with which the auction will suffer a large revenue loss. For any distribution $\pi \in \mathcal{H}$, let $\mathcal{B}_\pi$ denote the set of $i \in \{1,2,\dots,K-2\}$ for which algorithm A performs badly in the sense that the auction $A(\pi)$ suffers from an expected revenue loss more than $\frac{25\epsilon}{\eta K}$, conditioned on that the bidder's valuation type is i, with probability at least $\frac{3}{10}$:

$$\mathcal{B}_\pi = \left\{i:\ \Pr_{A(\pi)}\left[\Delta_{A(\pi),\pi}(i) > \frac{25\epsilon}{\eta K}\right] > \frac{3}{10}\right\}.$$

Lemma 12. *There exists $\pi^* \in \mathcal{H}$ such that:*

$$|\mathcal{B}_{\pi^*}| \geq \frac{K}{2} - 1.$$

In the rest of the analysis, we will focus on the distribution $\pi^* \in \mathcal{H}$ for which the conclusion of the above lemma holds. The above lemma is already good enough for proving a weaker claim that the expected revenue loss is $\Omega(\epsilon)$. However, Lemma 9 claims a stronger statement that the $\Omega(\epsilon)$ expected revenue loss must happen with a constant probability. To prove this, we need to further discuss the number of $i \in \{1, 2, \ldots, K-2\}$ for which the realized auction $A(\pi^*)$ performs poorly.

For any realization of the auction $A(\pi^*)$, let $\mathcal{B}^*{}_{A(\pi^*)}$ denote the set of $i \in \{1, 2, \ldots, K-2\}$ for which the returned auction $A(\pi^*)$ performs poorly in the sense that it suffers from a revenue loss more than $\frac{25\epsilon}{\eta K}$ conditioned on that the bidder's valuation type is i:

$$\mathcal{B}^*{}_{A(\pi^*)} = \left\{ i : \Delta_{A(\pi^*),\pi^*}(i) > \frac{25\epsilon}{\eta K} \right\}.$$

Lemma 13. *For the distribution $\pi^* \in \mathcal{H}$ in Lemma 12, with probability at least 0.01, we have:*

$$\left|\mathcal{B}^*{}_{A(\pi^*)}\right| \geq \frac{K}{25}.$$

Finally, we complete the proof of the Lemma 9 by arguing that the algorithm A must suffer from revenue loss at least ϵ on the distribution π^*. More specifically, when the conclusion of Lemma 13 is true, which happens with probability at least 0.01, we have the following sequence of inequalities:

$$\begin{aligned}\textsc{Opt}(\pi^*) - \textsc{Rev}(A(\pi^*), \pi^*) &= \sum_{\theta \in \Theta} \Delta_{A(\pi^*),\pi^*}(\theta) \cdot \pi(\theta) \\ \geq \eta \sum_{i \in \mathcal{B}^*_{A(\pi^*)}} \Delta_{A(\pi^*),\pi^*}(i) &> \frac{25\epsilon}{K} \cdot |\mathcal{B}^*_{A(\pi^*)}| \geq \epsilon.\end{aligned}$$

And with Lemma 9 we can now prove Theorem 2.

Proof of Theorem 2. With Lemma 9, Theorem 2 follows from straightforward calculations. We can upper bound $D_{\mathsf{SKL}}(\pi^1, \pi^2)$ (for any choice of i and $S' \subseteq \{1, \ldots, K-2\}/\{i\}$) by

$$\begin{aligned}&D_{\mathsf{SKL}}(\pi^1, \pi^2) \\ =&\eta \cdot \left(\sum_{\omega \in \Omega} \pi^1(\omega|i) \ln \frac{\pi^1(\omega|i)}{\pi^2(\omega|i)} + \sum_{\omega \in \Omega} \pi^2(\omega|i) \ln \frac{\pi^2(\omega|i)}{\pi^1(\omega|i)} \right) \\ =&O(K^{-1}\eta\alpha^2\epsilon^2),\end{aligned}$$

which is exactly the inverse of the lower bound we claim in the theorem. ☐

5. Generalization to the Multi-Bidder Case

In this section, we show how our sample complexity results can be generalized to the multi-bidder case.

Let there be n bidders. Without loss of generality, we assume that all bidders share a common valuation type set $\Theta = \{1, 2, \ldots, K\}$. Otherwise, we can take Θ as the union of each bidder's type set. For each bidder $i \in [n]$, let the external signal of the bidder be the joint types of other bidders $\boldsymbol{\theta_{-i}} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_n)$. We denote by α_i the degree of correlation between bidder i and other bidders and let $\alpha = \min_{i \in [n]} \alpha_i$.

5.1. Crémer-McLean Auction in n-Bidder Case

Consider the following Vickrey auction (x^*, p^*). For each bidder i, if she has the highest valuation among all bidders, $x_i^*(\boldsymbol{\theta}) = 1$ and $p_i^*(\boldsymbol{\theta})$ is the second-highest valuation. Otherwise, $x_i^*(\boldsymbol{\theta})$ and $p_i^*(\boldsymbol{\theta})$ are both zero. We denote by $u_i(\theta_i)$ the expected utility of bidder i in a Vickrey auction. That is,

$$u_i(\theta_i) = \sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} \pi(\boldsymbol{\theta}_{-i}|\theta_i) \left[v(\theta_i) x_i^*(\theta_i, \boldsymbol{\theta}_{-i}) - p_i^*(\theta_i, \boldsymbol{\theta}_{-i}) \right].$$

For each bidder i, under the Crémer-McLean condition, there exists a vector $\boldsymbol{p}_i = (p_i(\boldsymbol{\theta}_{-i}))_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}}$ such that for each $\theta_i \in \Theta$,

$$\sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} \pi(\boldsymbol{\theta}_{-i}|\theta_i) p_i(\boldsymbol{\theta}_{-i}) = u_i(\theta_i).$$

Then Crémer-McLean auction (x^*, p^{CM}) allocates the item to the bidder with highest valuation and charges each bidder a payment of $p_i^{\mathrm{CM}}(\theta_i, \boldsymbol{\theta}_{-i}) = p_i^*(\theta_i, \boldsymbol{\theta}_{-i}) + p_i(\boldsymbol{\theta}_{-i})$.

In short, Crémer-McLean auction in the multi-bidder case is equivalent to a Vickrey auction with additional payments. The additional payments fully extract the bidders' expected utilities in a Vickrey auction.

5.2. The Upper Bound

We present Algorithm 3 to learn a near-optimal n-bidder Crémer-McLean auction. The algorithm achieves the sample complexity bound as shown below.

Theorem 3. *For any $0 < \epsilon < 1$ and any α-strongly correlated distribution π with marginal probabilities at least η, Algorithm 3 returns an n-bidder auction with expected revenue at least $(1-\epsilon)$ of the full surplus with probability at least $1-\delta$, if the number of samples m satisfies*

$$m \geq 250 n^2 K^2 \eta^{-3} \alpha^{-2} \epsilon^{-2} \max\{5 \ln(12 n K \delta^{-1}), 8K\}. \tag{8}$$

The main idea of Algorithm 3 is to reduce the problem of learning an n-bidder Crémer-McLean auction to n problem

Algorithm 3 Multi-Bidder Empirical Crémer-McLean Auction

Input: m i.i.d. samples from the joint distribution π
Output: an auction that decides the allocation and payment
Step 1. Let $\hat{\pi}$ be the empirical joint distribution.
Step 2. For each bidder i and every $\theta_i \in \Theta$, let

$$\hat{u}_i(\theta_i) = \sum_{\boldsymbol{\theta}_{-i} \in \Omega_i} \hat{\pi}(\boldsymbol{\theta}_{-i}|\theta_i)[v(\theta_i)x_i^*(\theta_i, \boldsymbol{\theta}_{-i}) - p_i^*(\theta_i, \boldsymbol{\theta}_{-i})],$$

where $\hat{\pi}(\boldsymbol{\theta}_{-i}|\theta)$ represents $\hat{\pi}(\boldsymbol{\theta})/\sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} \hat{\pi}(\boldsymbol{\theta}_{-i}, \theta)$.
Step 3. Let

$$\hat{u}_i'(\theta_i) = \max\left\{0,\ \hat{u}_i(\theta_i) - \frac{5Kv(K)\zeta}{2\alpha}\right\}, \quad (9)$$

where $\|\hat{\boldsymbol{u}}_i\|$ is the Euclidean norm of the vector $\hat{\boldsymbol{u}}_i = (\hat{u}_i(1), \hat{u}_i(2), \ldots, \hat{u}_i(K))$ and

$$\zeta = \sqrt{\frac{10}{\eta m}} \max\left\{\sqrt{5\ln(12nK\delta^{-1})}, \sqrt{8K}\right\}.$$

Step 4. For each bidder i, run Algorithm 1 with the empirical distribution $\hat{\pi}$ and deflated expected utilities $\hat{u}_i'(\theta_i)$ as input to get the addition payment $\hat{p}_i(\boldsymbol{\theta}_{-i})$.
Step 5. Run a Vicrey auction and charge each bidder an additional payment of $\hat{p}_i(\boldsymbol{\theta}_{-i})$.

instances of learning bidder i's additional payment $p_i(\boldsymbol{\theta}_{-i})$. We apply Algorithm 1 to learn $p_i(\boldsymbol{\theta}_{-i})$ of each bidder i by simply replacing the down-shifted valuation $v'(\theta)$ with the deflated expected utility $\hat{u}_i'(\theta_i)$.

In Algorithm 3, there are two sources of errors which may lead to the additional payments of the bidders failing to fully extract the expected utilities. The first is the inaccurate estimation of the conditional probabilities $\pi(\boldsymbol{\theta}_{-i}|\theta)$ ($\boldsymbol{\theta}_{-i} \in \Theta'_{-i}$) on the left-hand side of Eq. 4, which we have analyzed in Section 3. The second source of revenue loss comes from the inaccurate estimation of the expected utilities $u_i(\theta_i)$ on the right-hand side of Eq. 4. We bound this error with the following lemma.

Lemma 14. *If the number of samples m satisfies Eq. 8, we have that*

$$|u_i(\theta_i) - \hat{u}_i(\theta_i)| \leq v(K)\zeta$$

holds for all bidder i and all $\theta_i \in \Theta$ with probability at least $1 - \frac{\delta}{2}$.

As the errors from both sources are bounded, we can prove interim IR and $(1-\epsilon)$ approximation of the auction returned by Algorithm 3, and thus complete the proof of Theorem 3.

Looking at Theorem 3, what might be surprising is that the sample complexity of Crémer-McLean auction does not depend on the joint type space of size K^n. This is because it is possible to learn a near-optimal auction even when we do not know each joint type's probability precisely. Take the estimation of $u_i(\theta_i)$ for an example. For bidder i with type θ_i, there are only θ_i possible values of $v(\theta_i)x_i^*(\theta_i, \boldsymbol{\theta}_{-i}) - p_i^*(\theta_i, \boldsymbol{\theta}_{-i})$. We divide the external signal set Θ_{-i} into θ_i subsets $\Omega_1, \ldots, \Omega_{\theta_i}$ corresponding to the highest valuation by others from $v(1)$ to no less than $v(\theta_i)$. To estimate $u_i(\theta_i)$, we only need to learn at most K conditional probabilities $\Pr[\boldsymbol{\theta}_{-i} \in \Omega_j|\ \theta_i]$ ($j \in \{1, 2, \ldots, \theta_i\}$), which takes much fewer samples than what is needed for estimating all of $|\Theta_{-i}|$ conditional probabilities $\pi(\boldsymbol{\theta}_{-i}|\theta_i)$ precisely.

On the other hand, the computational complexity of Algorithm 3 is polynomial in the size of the joint type space K^n. Note that this is inevitable due to the exponentially large representation of the joint distribution.

5.3. The Lower Bound

Our hard instances constructed in Section 4 can be easily rewritten as a family of 2-bidder joint distributions by letting the external signal of a bidder be the other bidder's type. We set the valuation function of bidder 1 to be the same as in Eq. 6 and let the valuations of bidder 2's types be $o(\epsilon)$. The smallest marginal probability of any distribution in the set $\mathcal{H}$ is $\Omega(\eta)$, and the singular values of the matrices Γ_1 and Γ_2 defined in Eq. 1 are at least α (see the Supplementary Material for reason). Then our lower bound of $\Omega(K\eta^{-1}\alpha^{-2}\epsilon^{-2})$ and its analysis in Section 4 naturally hold.

6. Conclusion

In this work we investigate the sample complexity of the optimal Crémer-McLean auction with correlated prior distributions. We present upper and lower bounds on the number of samples required to learn a auction that can extract a near-optimal revenue. These results provide us new insights on what could be the key factors of the prior distributions that decide the learnability of a near-optimal auction. In particular, it suggests that learning a near-optimal auction with correlated prior distributions is hard when there are many valuation types, or the level of correlation is low, or there is some valuation type with very small marginal probability. Conceptually, these results provide new evidence to the common belief that the Crémer-McLean auction is usually "too good to be true".

There are many directions in which this work could be extended. The first question is to close the gap between our upper and lower bounds. In addition, it is also interesting to consider the sample complexity of other families of auctions with correlated distributions, such as the lookahead auction (Ronen, 2001) and the correlated-robust auction (Bei et al., 2019).

References

Albert, M., Conitzer, V., and Lopomo, G. Assessing the robustness of cremer-mclean with automated mechanism design. In *AAAI*, pp. 763–769. Citeseer, 2015.

Albert, M., Conitzer, V., and Lopomo, G. Maximizing revenue with limited correlation: The cost of ex-post incentive compatibility. In *AAAI*, pp. 383–389, 2016.

Albert, M., Conitzer, V., Lopomo, G., and Stone, P. Mechanism design with correlated valuations: Efficient methods for revenue maximization. *Under Submission at Operations Research*, 2017a.

Albert, M., Conitzer, V., and Stone, P. Mechanism design with unknown correlated distributions: Can we learn optimal mechanisms? In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pp. 69–77, 2017b.

Barelli, P. On the genericity of full surplus extraction in mechanism design. *Journal of Economic Theory*, 144(3): 1320–1332, 2009.

Barrodale, I. and Stuart, G. Algorithm 576: A fortran program for solving ax= b [f4]. *ACM Transactions on Mathematical Software (TOMS)*, 7(3):391–397, 1981.

Bei, X., Gravin, N., Lu, P., and Tang, Z. G. Correlation-robust analysis of single item auction. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 193–208. SIAM, 2019.

Brustle, J., Cai, Y., and Daskalakis, C. Multi-item mechanisms without item-independence: Learnability via robustness. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pp. 715–761, 2020.

Chen, Y.-C. and Xiong, S. Genericity and robustness of full surplus extraction. *Econometrica*, 81(2):825–847, 2013.

Cole, R. and Roughgarden, T. The sample complexity of revenue maximization. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pp. 243–252, 2014.

Crémer, J. and McLean, R. P. Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica*, 53:345–361, 1985.

Crémer, J. and McLean, R. P. Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica: Journal of the Econometric Society*, pp. 1247–1257, 1988.

Devanur, N. R., Huang, Z., and Psomas, C.-A. The sample complexity of auctions with side information. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pp. 426–439, 2016.

Devroye, L. et al. The equivalence of weak, strong and complete convergence in l_1 for kernel density estimates. *The Annals of Statistics*, 11(3):896–904, 1983.

Fu, H., Haghpanah, N., Hartline, J., and Kleinberg, R. Optimal auctions for correlated buyers with sampling. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pp. 23–36, 2014.

Gonczarowski, Y. A. and Nisan, N. Efficient empirical revenue maximization in single-parameter auction environments. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pp. 856–868, 2017.

Guo, C., Huang, Z., and Zhang, X. Settling the sample complexity of single-parameter revenue maximization. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pp. 662–673, 2019.

Hartline, J. and Taggart, S. Sample complexity for non-truthful mechanisms. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pp. 399–416, 2019.

Heifetz, A. and Neeman, Z. On the generic (im) possibility of full surplus extraction in mechanism design. *Econometrica*, 74(1):213–233, 2006.

Huang, Z., Mansour, Y., and Roughgarden, T. Making the most of your samples. *SIAM Journal on Computing*, 47 (3):651–674, 2018.

Jin, Y., Lu, P., and Xiao, T. Learning reserve prices in second-price auctions. *arXiv preprint arXiv:1912.10069*, 2019.

Laffont, J.-J. and Martimort, D. Mechanism design with collusion and correlation. *Econometrica*, 68(2):309–342, 2000.

McAfee, R. P. and Reny, P. J. Correlated information and mecanism design. *Econometrica: Journal of the Econometric Society*, pp. 395–421, 1992.

Mitzenmacher, M. and Upfal, E. *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press, 2017.

Morgenstern, J. H. and Roughgarden, T. On the pseudo-dimension of nearly optimal auctions. In *Advances in Neural Information Processing Systems*, pp. 136–144, 2015.

Myerson, R. B. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

Peters, M. Surplus extraction and competition. *The Review of Economic Studies*, 68(3):613–631, 2001.

Robert, J. Continuity in auction design. *Journal of Economic theory*, 55(1):169–179, 1991.

Ronen, A. On approximating optimal auctions. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pp. 11–17, 2001.

Schechter, E. *Handbook of Analysis and its Foundations*. Academic Press, 1996.

Syrgkanis, V. A sample complexity measure with applications to learning optimal auctions. In *Advances in Neural Information Processing Systems*, pp. 5352–5359, 2017.