Targeting and Signaling in Ad Auctions

Ashwinkumar Badanidiyuru*   Kshipra Bhawalkar†   Haifeng Xu‡

Abstract

Modern ad auctions allow advertisers to target more specific segments of the user population. Unfortunately, this is not always in the best interest of the ad platform – partially hiding some information could be more beneficial for the platform's revenue. In this paper, we examine the following basic question in the context of second-price ad auctions: how should an ad platform optimally reveal information about the ad opportunity to the advertisers in order to maximize revenue? We consider a model in which bidders' valuations depend on a random state of the ad opportunity. Different from previous work, we focus on a more practical, and challenging, situation where the space of possible realizations of ad opportunities is *extremely large*. We thus focus on developing algorithms whose running time is polynomial in the number of bidders, but is *independent* of the number of ad opportunity realizations.

We assume that the auctioneer can commit to a *signaling scheme* to reveal noisy information about the realized state of the ad opportunity, and examine the auctioneer's algorithmic question of designing the optimal signaling scheme. We first consider that the auctioneer is restricted to send a *public* signal to all bidders. As a warm-up, we start with a basic (though less realistic) setting in which the auctioneer knows the bidders' valuations, and show that an ϵ-optimal scheme can be implemented in time polynomial in the number of bidders and $1/\epsilon$. We then move to a well-motivated Bayesian valuation setting in which the auctioneer and bidders both have private information, and present two results. First, we exhibit a characterization result regarding approximately optimal schemes and prove that any constant-approximate public signaling scheme must use exponentially many signals. Second, we present a "simple" public signaling scheme that serves as a constant approximation under mild assumptions.

Finally, we initiate an exploration on the power of being able to send different signals *privately* to different bidders. In the basic setting where the auctioneer knows bidders' valuations, we exhibit a polynomial-time private scheme that extracts *almost full surplus* even in the worst Bayes Nash equilibrium. This illustrates the surprising power of private signaling schemes in extracting revenue.

1 Introduction

Every single day, through various online advertising platforms, advertisers' ads are seen by billions of web users. Each ad opportunity, usually associated with a randomly arrived web user, can be depicted by various characteristics of the web user. A particular advantage of online advertising is that it can allow advertisers to target specific user populations. To do so, the platform (i.e., the *auctioneer*) can selectively reveal some characteristics of the web user to advertisers (i.e., *bidders*). For example, the function of allowing advertisers to set different keywords and bids for different ad campaigns, as permitted by most ad platforms, is essentially equivalent to revealing partial information to advertisers before their bidding. Features like these have motivated the following basic question in ad auctions: how much information regarding the web user should the platform reveal to advertisers? This is particularly important given increasingly refined information about web users known by today's ad platforms like Facebook or Google. In this paper, we examine this question under the objective of maximizing the auctioneer's *revenue*.

Like [11, 5], we focus on the single-item second-price auction and assume that each advertiser's valuation for the ad opportunity depends on a common random *state of nature*. A natural interpretation of the state of nature is the characterization of the randomly arrived web user, which affects advertisers' valuations for the ad opportunity. The state of nature is drawn from a common-knowledge prior distribution, however only the auctioneer can observe its realization. We model the auctioneer's information revelation policy as a *signaling scheme*, which is a randomized map from the realized state of nature to a set of signals. There are two natural signaling models: *public* and *private* signaling schemes. In a public signaling scheme, the auctioneer is constrained to send a public signal to bidders, therefore each bidder performs a Bayesian update and reaches a common posterior belief regarding the state of nature. When the auctioneer is allowed to send different signals to different bidders, the signaling scheme is private.

We consider both the *known valuation* and *Bayesian valuation* settings. In the known valuation setting, the auctioneer knows bidders' values at any given realized state. Though admittedly less realistic, the known valuation setting is a natural first step for studying signaling in auctions and has been examined by several previous works [11, 5, 17]. In the (more realistic) Bayesian valuation setting, we assume that each bidder's value depends on a random private bidder type that is unknown to the auctioneer. To account for bidders' strategic behaviors in the auction, we adopt the

---
*Google Research, ashwinkumarbv@google.com.

†Google Research, kshipra@google.com.

‡University of Southern California, haifengx@usc.edu. Part of this work was conducted while Haifeng Xu was an intern at Google Research, Mountain View, CA.

dominant-strategy truth-telling equilibrium whenever it exists, and otherwise, the worst Bayes Nash equilibrium (particularly, in private signaling schemes). We remark that the signaling schemes considered in [11, 5] are all public.

When designing signaling schemes for modern ad auctions, a primary computational challenge is the large diversity of web users that an ad opportunity might correspond to. Web users vary in their interests, demographics, browsing history, etc. In the advertising system, such user information is usually captured by thousands of features and each of their (exponentially many) configurations could possibly be realized. Translated to the model, this means that the space of possible realizations of the state of nature can be *extremely large*. One key difference between our work and previous work on signaling in auctions is that we take into account this challenge of extremely large number of states of nature, and restrict our attention to algorithms with running time *independent* of this parameter.

We note that signaling naturally relates to *targeting* in ad auctions. In fact, as we will elaborate later, the Bayesian valuation setting we consider is motivated by the ad auction feature called *re-targeting*. A signaling scheme essentially specifies the extent to which each advertiser can target the web users. Unsurprisingly, it is usually not in the best interest of the auctioneer to reveal full information to advertisers: over refined targeting can result in fewer advertisers competing for the ad opportunity and possible a large gap between the highest and second-highest bid [17, 19]. This is detrimental to the revenue in second price auctions. On the other hand, completely obscuring user information is also suboptimal since advertisers will bid conservatively due to their uncertainty regarding the value of the ad opportunity. Our goal is thus to design optimal signaling schemes that properly balance the amount of competition and bidders' valuations in the auction. Next, we elaborate the results in detail.[1]

1.1 Our Models and Results. We first consider the design of optimal public signaling schemes. As a warm-up, we start with the known-valuation setting and show that a simple Monte-Carlo algorithm implements an ϵ-optimal public scheme and runs in $\text{poly}(n, \frac{1}{\epsilon})$ time where n is the number of bidders. Note that this setting is also studied by [11, 5], however the running time of their algorithms depends polynomially on the number of states of nature.

Next, we move to a natural Bayesian-valuation model motivated by *re-targeting*. Re-targeting is a feature in ad auctions that allows each advertiser to target a set of web users (e.g., those who visited the advertiser's website) and use a different bid for users from that set.[2] In this setting, the state of nature is a *vector* of binary feature $\theta_i \in \{0, 1\}$ for each bidder i, indicating whether the web user is the bidder's targeted ones or not. Therefore, the number of possible states is exponential in the number of bidders. We assume that bidder i's valuation is determined by θ_i and a private bidder type that is i.i.d. across bidders. Note however, θ_i's may be *correlated* since some bidders may target similar items. For this Bayesian-valuation model, we first prove that any public signaling scheme that uses sub-exponentially many signals can *not* guarantee any constant fraction of the optimal revenue. We note that, though *not* a measure of computational complexity, the number of signals can be viewed as a natural complexity measure of the signaling scheme, which is analogous to the menu-size complexity studied in the optimal mechanism design literature [15, 29, 2]. Complementing this result, we show a signaling scheme that serves as a constant approximation under mild assumptions. Our scheme is "simple" to implement and tracks the following conventional wisdom. In particular, it reveals full information to bidders whenever there is enough competition in the auction and otherwise, it pools bidders to form competition by selectively obscuring information.

Finally, we initiate an exploration on the power of private signaling schemes, which to the best of our knowledge has not been studied in the context of auctions. We first exhibit an example showing that the optimal private scheme can outperform the optimal public scheme *arbitrarily* in terms of the auctioneer's revenue. This motivates the algorithmic question of computing the optimal private signaling scheme. Here, we restrict our attention to the known-valuation setting, which is basic but already reveals sufficient nontriviality due to intricate equilibrium bidding behaviors and the equilibrium selection issue. We present a polynomial-time private signaling scheme that extracts almost the *full surplus* even in the *worst* Bayes Nash equilibrium. This illustrates the surprising power of pri-

---

[1]Note that though our work is primarily motivated by ad auctions, our results are applicable to any second-price auction with large number of possible item realizations.

[2]Re-targeting is allowed in several ad platforms including Google's and Facebook's. Each advertiser will specify a targeted set of web users (usually using their user IDs) in the platform system. When a web user arrives, the platform informs each bidder whether the user is in their targeted set or not. Then each bidder submits a bid. We refer curious readers to the help pages from various companies for more details, e.g., `https://support.google.com/adwords/answer/1704368?hl=en`.

vate signaling schemes in extracting revenue.

1.2 Additional Related Work.

1.2 Additional Related Work. The study of information asymmetry in auctions dates back as early as 1980s [25, 22, 20, 21]. In a seminal paper, Milgrom and Weber [21] prove that in common-value auctions with bidders of symmetric random signals, among all public signaling schemes, revealing full information maximizes the auctioneer's revenue in most commonly seen auction formats (i.e., descending, ascending, second-price and first-price auctions). However, later work has observed that such full transparency is not optimal in general, and the auctioneer could increase revenue by revealing partial information to bidders [23, 6, 13, 26]. This has motivated a surge of recent interest in the algorithmic question of computing the optimal signaling scheme that maximizes revenue in a given auction format. [14] prove that revealing full information is optimal, among all public schemes, in Myerson's optimal auction. However, the optimal scheme is usually a non-trivial one and involves randomization in second price auctions [11, 5]. Recently, [8] characterize the optimal joint design of the signaling scheme and auction mechanism.

To our knowledge, [11, 5] are the first to study optimal signaling design in the known-valuation setting. They observe a linear program for computing the optimal public scheme in time polynomial in the number of bidders and states of nature. However, this is not satisfactory for ad auctions since the space of all possible states of nature is usually extremely large. It is generally impractical to solve such a large LP quickly. Though not explicitly stated, [17] also study the known-valuation setting with exponentially many states of nature. However, they do not consider the design of optimal signaling but instead compare the revenue between two simple schemes, i.e., revealing no information or full information.

[11] also study the Bayesian valuation setting, though in a more idealized model. Specifically, they assume that given a state of nature, the auctioneer's uncertainty regarding the bidders's valuation is captured by a distribution over polynomially many value profiles. This is somewhat restrictive since the space of value profiles typically increases exponentially in the number of bidders. However, even under this simplification, it is shown that computing the optimal public signaling scheme is NP-hard. In a follow-up work, [7] derive a PTAS for the Bayesian model of [11]. Their algorithm crucially relies on the fact that there are polynomially many states of nature, thus does not apply to our setting.

All the mentioned algorithmic work so far focuses on public signaling. Recent work of [1, 3] studies optimal private signaling for agents with no inter-dependence. Therefore, given a signal, each agent makes a decision that is independent of others. The problem usually becomes more intricate when there are strategic interactions among agents (e.g., in an auction). [4] characterize the distributions of players' action profiles at equilibrium under an asymmetric information structure as the set of Bayes correlated equilibria. To our knowledge, there has not been much progress on the design of optimal private signaling in such strategic settings. [27] characterizes the optimal private signaling in a two-agent two-action game. [28] consider private signaling in non-atomic routing, and focus on analyzing the cost decrease via private signaling in a graph of parallel links.

2 Preliminaries

2.1 The Model. We focus on the single-item second price auction, which is widely adopted by ad exchange platforms for selling display ads, video ads, etc. Let n denote the total number of bidders ("he" or "them"). Vector $v = v(\theta) \in \mathbb{R}^n_+$ denotes the profile of bidders' valuations which depends on a random *state of nature* θ. We assume that the state of nature is a-priori unknown to the bidders, and drawn from a common-knowledge prior distribution λ supported on set Θ. The auctioneer ("she"), on the other hand, can observe the realized θ. For ease of presentation, we assume Θ is finite, though this is not crucial for our results. In ad auctions, θ could encode the information of a web page viewer (e.g., demographic information, browsing history, etc.), which is known to the auctioneer. We call this the **Known-Valuation Setting** (KVS) since in this case, the auctioneer is certain about each bidder's value after observing θ. We use $v_i(\theta)$ to denote bidder i's value and $\mathcal{V}_i = \{v_i(\theta) : \theta \in \Theta\}$ is its support. Naturally, $v(\theta) \in \mathcal{V} = \prod_{i=1}^n \mathcal{V}_i$ for any θ. We assume the set Θ to be *extremely large*, and will focus on algorithms with running time independent of $|\Theta|$. However, any algorithm can query the distribution λ for the probability λ_θ of a state θ and can also sample from λ.

We also study the **Bayesian-Valuation Setting** (BVS), in which bidders' values have two-sided uncertainties. Motivated by the *re-targeting* feature in ad auctions, we focus on a basic model in which the state of nature consists of a *binary* feature $\theta_i \in \{0, 1\}$ for each bidder i, indicating whether the item is among bidder i's targeted ones or not. Therefore, $\theta = (\theta_1, ..., \theta_n) \in \{0, 1\}^n = \Theta$. We assume bidder i's value $v_i = v(\theta_i, t_i)$ is a function of θ_i and a private bidder type t_i, which bidder i is privy to, but is unknown to the auctioneer. Instead, the auctioneer only knows the prior distribution of t_i. We assume that the bidder type t_i's are inde-

pendent and identically distributed (i.i.d.). Therefore, $v_i = v(1, t_i)$ follows the same distribution for each i, denoted as $H(v)$, and $v_i = v(0, t_i)$ follows another distribution $L(v)$. We also call $H[L]$ the *high* [*low*] distribution for convenience, since it corresponds to the value distribution of targeted [un-targeted] items. W.l.o.g., assume $\mathbf{E}_{v \sim H}[v] > \mathbf{E}_{v \sim L}[v]$. Notice however, we allow the binary feature θ_i's to be correlated since some bidders may target similar items. If, moreover, $\{\theta_i\}_{i \in [n]}$ are also i.i.d., we say the bidders are *symmetric*.

2.2 Signaling Schemes. We assume that the auctioneer can commit to a *signaling scheme* that reveals noisy information about the realized θ. A signaling scheme φ is a *randomized* map from the states in Θ to a set of *signals* Σ. We use $\varphi(\theta, \sigma)$ to denote the probability of selecting the signal $\sigma \in \Sigma$ given the state θ. Notice that φ is public knowledge. An algorithm implements signaling scheme φ if it takes as input a state of nature θ, and outputs σ with probability $\varphi(\theta, \sigma)$. Our goal is to design efficient and (approximately) optimal *revenue-maximizing* signaling schemes. When φ yields expected revenue within an additive [multiplicative] ϵ of the best possible, we say it is ϵ-optimal [ϵ-approximate] in the additive [multiplicative] sense. We consider two natural signaling models.

Public Signaling Schemes. In a *public* signaling scheme, the auctioneer is only allowed to send a public signal, so that each bidder infers the same information. Let $\alpha_\sigma = \sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma)$ denote the probability of sending the public signal σ. Upon receiving σ, bidders infer a common posterior distribution of θ by Bayesian updating. In particular, $\mathbf{Pr}(\theta | \sigma) = \lambda_\theta \varphi(\theta, \sigma) / \alpha_\sigma$. In the known-valuation setting, we use

$$v(\sigma) = \mathbf{E}_{\theta | \sigma}[v(\theta)] = \frac{1}{\alpha_\sigma} \sum_\theta \lambda_\theta \varphi(\theta, \sigma) \cdot v(\theta)$$

to denote the expected value profile conditioned on signal σ, and $v_i(\sigma)$ is bidder i's conditional expected value. Observe that, conditioned on receiving the public signal σ, bidding $v_i(\sigma)$ is the dominant strategy for each bidder i (see, e.g., [11]) in the second price auction. Our goal is to maximize the auctioneer's expected revenue at the (unique) dominant-strategy truth-telling equilibrium, expressed as follows in the known-valuation setting:

$$\text{KVS: } \mathbf{Rev}(\varphi) = \sum_{\sigma \in \Sigma} \alpha_\sigma \cdot \max_i 2 [v_i(\sigma)]$$
$$= \sum_{\sigma \in \Sigma} \max_i 2 \left[\sum_\theta \lambda_\theta \varphi(\theta, \sigma) \cdot v_i(\theta) \right]$$

where the function max2 returns the second largest value of a given set or vector. The following lemma

shows that an optimal public signaling does not need to use more than $n(n-1)$ signals in the known-valuation setting. The proof follows a simple revelation-principal-type argument.

LEMMA 2.1. *(Adapted from [18, 11]) There always exists an optimal* public *signaling scheme in the* known-valuation *setting that uses at most $n(n-1)$ signals, with signal σ^{ij} ($i \neq j \in [n]$) resulting in bidder i [j] as the bidder with the highest [second-highest] value.*

In the Bayesian-valuation setting, the description of a public signaling scheme remains the same. Therefore, the expected revenue is expressed similarly, but with an additional expectation over the randomness of bidder's private types.

$$\text{BVS: } \mathbf{Rev}(\varphi) = \mathbf{E}_{t_1, \dots, t_n} \left[\sum_{\sigma \in \Sigma} \max_i 2 \left[\sum_\theta \lambda_\theta \varphi(\theta, \sigma) \cdot v(\theta_i, t_i) \right] \right]$$

We remark that Lemma 2.1 does not hold in the Bayesian-valuation setting. Actually, in sharp contrast to Lemma 2.1, we prove in Section 4 that in BVS, any public signaling scheme using sub-exponentially many signals cannot guarantee even a constant fraction of the optimal revenue.

Note that among all public signaling schemes, the one that maximizes the *social welfare* is to reveal full information to bidders, in both known-valuation and Bayesian-valuation settings [10]. Moreover, the optimal welfare is always an upper bound for the optimal revenue of public signaling. We remark that though the signaling model assumes that each bidder knows the prior distribution and signaling scheme, in practice bidder i only needs to know his conditional valuation $v_i(\sigma) = \mathbf{E}_{\theta | \sigma} v_i(\theta)$ on signal σ in order to figure out his optimal bid. $v_i(\sigma)$ can be estimated within good precision by samples (i.e., interacting with the system for enough times).

Private Signaling Schemes. In a private signaling scheme, the auctioneer is allowed to send different (possibly correlated) signals to different bidders. That is, bidders are asymmetrically informed. After receiving the private signal, each bidder updates his own belief regarding θ. Obviously, the public signaling scheme is a special class of private signaling schemes. We will restrict our attention to the *known-valuation setting* when studying private signaling schemes. An important observation regarding private signaling is that bidding the conditional expected value is generally not an equilibrium strategy for bidders. Moreover, the phenomenon of winner's curse may arise when bidders are asymmetrically informed, as illustrated in the following example.

EXAMPLE 1. (NON-TRUTHFULNESS AND WINNER'S CURSE IN PRIVATE SIGNALING) *Consider a second-*

price auction with 2 bidders in the known-valuation setting. There are two possible value profiles, listed in the following table together with corresponding probabilities.

	bidder 1	*bidder 2*
prob 0.5	1	2
prob 0.5	7	7

Consider the private signaling scheme that reveals the realized profile (i.e., full information) to bidder 2 but reveals no information to bidder 1. Since bidder 2 knows her value exactly, bidding the true value is the dominant strategy for bidder 2. Given no information, bidder 1 has an expected value of 4, which however is not her equilibrium bid. In fact, bidder 1 will not bid any value between 2 and 7 since in this case she can only win at the less-valued situation – the winner's curse. The equilibrium bid for bidder 1 turns out to be any bid under 2 and there are infinitely many equilibria under this private signaling scheme.

Example 1 reveals several challenges for computing the optimal private signaling scheme. The first is the equilibrium selection problem due to the existence of multiple equilibria, as illustrated by Example 1. There is usually not a natural equilibrium selection rule in private signaling. Second, even if we know which equilibrium to adopt, computing a desired equilibrium is non-trivial in second-price auctions with asymmetrically informed bidders. In fact, most prior work is only able to characterize equilibria under restrictive assumptions, e.g., two bidders, common valuation [16, 12, 26].

3 Warm-up: Public Signaling in the Known-Valuation Setting

We start with the design of the optimal public signaling scheme in the known-valuation setting, and prove the following theorem.

THEOREM 3.1. *Assume $v(\theta) \in [0,1]^n$ for all θ. An (additive) ϵ-optimal signaling scheme can be implemented in $\mathrm{poly}(n, \frac{1}{\epsilon})$ time. In particular, the running time is independent of $|\Theta|$.*

The proof of Theorem 3.1 combines and generalizes two ideas from previous literature. We present a sketch here, while defer the full proof to Appendix A. First, as observed in [11, 5], there is a $\mathrm{poly}(n, |\Theta|)$-sized linear program that explicitly computes the optimal public signaling scheme. Specifically, by Lemma 2.1, we can w.l.o.g. assume $\Sigma = \{\sigma^{ij}\}_{i \neq j \in [n]}$ in which σ^{ij} results in bidder i [j] to have the highest [second-highest] expected value. Then Linear Program (3.1) with variables $\varphi(\theta, \sigma)$ computes the optimal public signaling scheme.

However, any naive approach for solving LP (3.1) requires $\mathrm{poly}(n, |\Theta|)$ time, which is prohibitive when $|\Theta|$ is extremely large. To overcome this challenge, we employ a similar Monte-Carlo sampling algorithm as proposed by [9]. In particular, given as input a state of nature θ, we take $\mathrm{poly}(n, \frac{1}{\epsilon})$ many samples from λ, and let $\widetilde{\lambda}$ be the empirical distribution over these samples plus the given state θ. The algorithm then solves the linear program by instead using $\widetilde{\lambda}$ as the prior and relaxing the first two constraints of LP (3.1) by $1/\mathrm{poly}(1/\epsilon, n)$ (since the samples cannot exactly preserve these constraints), and signals for θ as suggested by the solution of the LP.

(3.1)

$$
\begin{aligned}
\text{maximize} \quad & \sum_{\sigma^{ij}} \sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma^{ij}) v_j(\theta) \\
\text{subject to} \quad & \sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma^{ij}) v_i(\theta) \geq \\
& \quad \sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma^{ij}) v_j(\theta), \quad \text{for } j \neq i. \\
& \sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma^{ij}) v_j(\theta) \geq \\
& \quad \sum_{\theta \in \Theta} \lambda_\theta \varphi(\theta, \sigma^{ij}) v_k(\theta), \quad \text{for } k \neq i, j. \\
& \sum_{\sigma^{ij} \in \Sigma} \varphi(\theta, \sigma) = 1, \quad \text{for } \theta. \\
& \varphi(\theta, \sigma^{ij}) \geq 0, \quad \text{for } \theta, \sigma^{ij}.
\end{aligned}
$$

We remark that though the idea for our Monte-Carlo algorithm is similar to that in [9], the analysis there does not directly apply to our setting. Particularly, the algorithm in [9] gives a *bi-criteria* FPTAS for the Bayesian persuasion problem (i.e., ϵ-loss in both optimality and incentive compatibility), and they prove that the bi-criteria loss is inevitable in their setting due to information-theoretic reasons. However, our algorithm is a *single-criteria* FPTAS and its analysis relies on the special property of the revenue objective.

4 Public Signaling in the Bayesian-Valuation Settings

In this section, we turn our attention to the more realistic Bayesian valuation setting, and consider the design of optimal public signaling schemes. We focus on a basic yet well-motivated model, i.e., $\theta = \{\theta_1, ..., \theta_n\} \in \{0, 1\}^n$ and bidder i's private information is captured by a type t_i which is i.i.d. across bidders. Note however, θ_i's can be correlated. When $\theta_i = 1$, bidder i's value $v_i = v(1, t_i)$ follows a common (high) distribution $H(v)$ for every i. Similarly, $v_i = v(0, t_i)$ follows the (low) distribution $L(v)$. We assume w.l.o.g $\mathbf{E}_{v \sim H}[v] > \mathbf{E}_{v \sim L}[v]$. If, moreover, θ_i's are i.i.d., we say that bidders are *symmetric*.

4.1 Lower Bounding the Number of Signals Required by Approximately Optimal Schemes.

We first bound the minimum number of signals required by any approximately optimal public signaling scheme. This is an important feature not only because it is

a basic property regarding the optimal scheme but also it serves as a natural complexity measure of the optimal scheme, which is analogous to the menu-size complexity studied in the recent mechanism design literature [15, 29, 2]. Several previous works have established polynomial-size *upper bounds* on the number of signals required by the optimal scheme in various settings [18, 11, 5]. In contrast, we prove that in our Bayesian-valuation setting, it is generally necessary to use *exponentially* many signals, even when only a constant-approximate public signaling scheme is sought.

THEOREM 4.1. *In the Bayesian-valuation setting with n symmetric bidders, any public signaling scheme using $2^{o(n^{1/4})}$ signals cannot guarantee any constant (multiplicative) fraction of the optimal revenue.*

We remark that Theorem 4.1 is *not* a computational hardness result, though it does indicate some challenges in designing the optimal public signaling scheme for the BVS setting. The proof of Theorem 4.1 is rather involved. We defer the full proof to Appendix B, and only present a sketch here.

The proof examines the following instance in the symmetric bidder setting: there are n bidders and each bidder's valuation follows either a (high) Bernoulli distribution with success probability $\frac{1}{\sqrt{n}}$ (when $\theta_i = 1$) or a (low) point distribution at 0 (when $\theta_i = 0$); θ_i's are i.i.d. with $\mathbf{Pr}(\theta_i = 1) = \epsilon$ for a small constant ϵ.

The proof is divided into two steps. We first show that the revenue of the optimal public signaling scheme is at least $1 - o(1)$. Intuitively, this is because when revealing full information, with high probability there are $\Omega(\epsilon n)$ bidders having the high value distribution and such competition results in a revenue close to 1. Second, we argue that the revenue of any public signaling scheme using $2^{o(n^{1/4})}$ signals must be far less than 1. The main technical challenge we overcome is in the second step: given an *arbitrary* public signaling scheme using at most N signals, we look to upper bound its revenue in terms of N.

In fact, instead of upper bounding the revenue, we upper bound the welfare of an arbitrary public signaling scheme using at most N signals since the revenue is then trivially upper bounded by the welfare. We first observe that among the public signaling schemes with at most N signals that achieve the maximum welfare, there is always a deterministic one that deterministically maps a state to a signal. Such a deterministic scheme can be viewed as a partition of $\Theta = 2^{[n]}$ into N subsets, denoted as $\{S_i\}_{i=1}^{N}$. Let $\mathbf{Pr}(S_i) = \sum_{\theta \in S_i} \lambda_\theta$ denote the total probability mass of the states in set S_i. We then prove the following key lemma (informal): for any S_i, either the welfare conditioned on S_i is $O(\epsilon)$ or $\mathbf{Pr}(S_i)$ is

exponentially small (specifically, $\mathbf{Pr}(S_i) \leq 2^{-\Theta(n^{1/4})}$). This lemma implies that any deterministic scheme with $2^{o(n^{1/4})}$ signals must have $O(\epsilon)$ welfare. The proof of the lemma involves intricate combinatorial analysis. At a high level, we view each state θ as a binary vector and let $\bar{\theta}$ denotes the expectation of the θ's in S_i. Note that each entry of $\bar{\theta}$ has expectation ϵ a-priori. We show that if m is the number of entries in $\bar{\theta}$ that are at least 2ϵ, then $\mathbf{Pr}(S_i) \leq e^{-2\sqrt{m}}$. Intuitively, this is because the sum of the expectation of these m entries is ϵm, so the probability that the sum is at least $2\epsilon m$ should be small. However, the concrete argument is more involved since we have to argue for an arbitrary deterministic set S_i while not a sampled one. Since the high distribution is a Bernoulli with success probability $1/\sqrt{n}$, if $m = o(\sqrt{n})$, the welfare from $\bar{\theta}$ would be $O(\epsilon)$; otherwise $m = \Omega(\sqrt{n})$ implying $\mathbf{Pr}(S_i) \leq 2^{-\Theta(n^{1/4})}$. This yields a proof sketch of the key lemma.

4.2 A Constant Approximate Scheme. In this section, we present a "simple" scheme that serves as a constant approximation under mild assumptions. We start by showing that revealing full information – which indeed uses exponentially many signals – can be arbitrarily far from optimality.

EXAMPLE 2. (FAILURE OF THE FULL INFORMATION SCHEME) *Consider the following simple example: the value of each bidder either follows the (high) uniform distribution in $[0, 1]$ (denoted as $U[0, 1]$) or follows the (low) point distribution at 0. The state of nature is a uniformly drawn bidder whose value follows $U[0, 1]$ while all other bidders have valuation 0. Only the auctioneer knows which bidder has a high value distribution. If the auctioneer reveals full information, the revenue is 0 since only one bidder has a non-zero valuation. On the other hand, consider the following signaling scheme: send a pooling signal $\sigma_{2i-1,2i}$ for $i \in 1, 2, ..., \lfloor n/2 \rfloor$ when either bidder $2i - 1$ or $2i$ has the high value distribution. That is, the scheme pools bidder $2i - 1$ and bidder $2i$ to create competition. Now from bidders' perspective, whenever the pooling signal $\sigma_{2i-1,2i}$ is sent, bidder $2i-1$ and $2i$ know that their valuations are drawn from $U([0, 1])$ with probability 0.5 and are 0 otherwise. As a result, bidder $2i - 1$ and $2i$ have expected value distribution $U([0, 0.5])$. Some calculation reveals that the expected revenue conditioned on signal $\sigma_{2i-1,2i}$ is $1/6$, which is arbitrarily better than 0.*

Revisiting Example 2, we note that the failure of the full information scheme there is due to the lack of competition. In particular, by revealing full information, the realized state includes only one non-zero-valued bidder, resulting in 0 revenue. We call these

the *tail states*, meaning the extreme situation with little competition in the auction. More specifically, we call the states in $\Theta_1 = \{\theta : |\theta| = 1\}$ the tail states. The high-revenue scheme described in Example 2 pools bidders to create competition at these tail states. Obviously, this is not always possible for any prior distribution λ. For example, if λ is a point distribution supported at the tail state $\theta = (1, 0, 0, ..., 0)$, then we cannot pool θ with any other tail states.

We thus first characterize when such tail-pooling is possible. Since the state of nature θ is in $\{0, 1\}^n$, for convenience we sometimes view θ as a *set of bidders* specified by entries of value 1. Therefore, each tail state can be treated as a bidder and Θ_1 can be treated as the set of all bidders $[n]$. We would like to pair the tail states in Θ_1 (possibly randomly) in a way such that each bidder in a pair has *equal* probability mass (like the scheme in Example 2). Such a *pooling scheme* can be captured by a randomized map $\pi : \Theta_1 \to \Theta_1$ which maps any given state $\theta \in \Theta_1$ to the state that it will be paired with. Let $\pi(\theta, \theta')$ denote the probability of mapping θ to θ'. The *feasible* pooling scheme π's are captured by the following linear system.
(4.2)
$$\begin{array}{ll} \sum_{\theta' \in \Theta_1} \pi(\theta, \theta') = 1, & \text{for } i \in \Theta_1. \\ \lambda_\theta \cdot \pi(\theta, \theta') = \lambda_{\theta'} \cdot \pi(\theta', \theta), & \text{for } \theta, \theta' \in \Theta_1. \\ \pi(\theta, \theta') \geq 0 \quad \text{and} \quad \pi(\theta, \theta) = 0, & \text{for } \theta, \theta' \in \Theta_1. \end{array}$$

Particularly, the second constraint guarantees that conditioned on that θ, θ' are pooled together, they will have equal probability mass. As pointed out earlier, linear system (4.2) is not always feasible and its feasibility depends on the portion of λ on set Θ_1. The following lemma shows that under a minor assumption on λ, linear system (4.2) is guaranteed to be feasible. The proof of Lemma 4.1 can be found in Appendix C.2.

LEMMA 4.1. *Linear system* (4.2) *is feasible if and only if for any* $\theta \in \Theta_1$, $\lambda_\theta \leq \sum_{\theta' \in \Theta_1 : \theta' \neq \theta} \lambda_{\theta'}$.

In other words, there exists a feasible pooling scheme if no state in Θ_1 has probability mass that is larger than the sum of *all other* states' probability masses. In this case, we say λ is tail-balanced. Particularly, λ is *tail-balanced* if $\lambda_\theta \leq \sum_{\theta' \in \Theta_1 : \theta' \neq \theta} \lambda_{\theta'}$ for any $\theta \in \Theta_1$. It is easy to see that a feasible pooling scheme (if exists) can be computed efficiently. Utilizing any feasible pooling scheme, we consider the signaling scheme that simply reveals θ *unless* $\theta \in \Theta_1$ in which case it pools θ with another $\theta' \in \Theta_1$ using a pooling scheme. The details are in Algorithm 1, which we term the *tail-pooling signaling scheme*.

The output of Algorithm 1 simply either reveals the true state (when $\sigma = \theta$) or reveals a set of two

---

Algorithm 1 Tail-Pooling Signaling Scheme for Bayesian Valuation Setting

---

Parameter: any feasible pooling scheme π.
Input: State of nature $\theta \in \{0, 1\}^n$.
Output: Signal σ
1: **if** $|\theta|_1 \geq 2$ or $|\theta|_1 = 0$ **then**
2:     $\sigma = \theta$;
3: **else**
4:     Pooling tail states: sample $\theta' \in \Theta_1$ with probability $\pi(\theta, \theta')$, and let $\sigma = \{\theta, \theta'\}$;
5: **end if**
6: **return** σ.

---

states (when $\sigma = \{\theta, \theta'\}$). Conditioned on $\sigma = \{\theta, \theta'\}$, the state θ and θ' show up both with probability $1/2$, as imposed by the second constraint in linear system (4.2). The following theorem shows that under mild assumptions, Algorithm 1 serves as a constant approximation to the optimal public signaling scheme.

THEOREM 4.2. *Assume the high and low distributions satisfy the monotone hazard rate (MHR) condition and λ is tail-balanced. When $n \geq 12$, Algorithm 1 achieves revenue that is at least $\frac{1}{8}$ fraction of the optimal welfare, thus is a $\frac{1}{8}$-approximate revenue-maximizing signaling scheme.*

Observe that the optimal welfare is obtained by revealing full information, thus is the expectation of the welfare at each state. Theorem 4.2 is proved by showing that for every state $\theta \in \Theta$, the revenue at θ in Algorithm 1 is at least $1/8$ fraction of the welfare at θ in the full information scheme. The following lemma summarizes the concrete bounds for different cases. Utilizing MHR properties, our proof establishes various inequalities comparing the welfare and revenue of a standard second price auction when each bidder's value is drawn independently from the high or low value distribution. The details are relegated to Appendix C.2.

LEMMA 4.2. *For any $\theta \in \Theta$, let $\mathbf{Rev}_{pool}(\theta)$ denote the revenue of the tail-pooling signaling scheme in Algorithm 1 at state of nature θ and $\mathbf{Wel}_{full}(\theta)$ denote the welfare of the full information scheme at state of nature θ. Then:*

1. *When $|\theta| = 0$ and $n \geq 2$, we have*

$$\mathbf{Rev}_{pool}(\theta) \geq \max\{\tfrac{1}{3}, \tfrac{\ln(n+1)-1}{\ln(n+1)}\} \cdot \mathbf{Wel}_{full}(\theta).$$

2. *When $|\theta| \geq 2$, we have*

$$\mathbf{Rev}_{pool}(\theta) \geq \max\{\tfrac{1}{6}, \tfrac{\ln(|\theta|+1)-1}{2\ln(|\theta|+1)}\} \cdot \mathbf{Wel}_{full}(\theta).$$

3. When $|\theta| = 1$ and $n \geq 12$, we have

$$\mathbf{Rev}_{pool}(\theta) \geq \tfrac{1}{8} \cdot \mathbf{Wel}_{full}(\theta).$$

REMARK 1. *Theorem 4.2 conveys an interesting conceptual message. That is, the simple "rule of thumb" principle, which reveals full information when there is enough competition and otherwise pools several bidders to form competition, is approximately optimal. As an instantiation of this principle, Algorithm 1 pools the tail states consisting of a single high-valuation bidder to form competition among two bidders. We make this choice because such a pooling scheme exists under only a minor assumption on λ, i.e., tail balance (see Lemma 4.1). Nevertheless, from a practical perspective, if there exist pooling schemes that can form competition among more bidders by pooling the states consisting of a few high-valuation bidders, we suspect that the analogy of Algorithm 1 may yield more revenue in practice.*

5 Private Signaling in the Known-Valuation Setting

In this section, we consider the design of the optimal private signaling scheme in the known-valuation setting. To see the advantage of using private schemes, we first exhibit an example showing that private signaling schemes may achieve *arbitrarily* better revenue than the optimal public signaling scheme.

EXAMPLE 3. *Consider a second-price auction with 3 bidders in the known-valuation setting. There are two value profiles, as listed in the following table with corresponding probabilities.*

	bidder 1	bidder 2	bidder 3
prob: $1 - \epsilon$	2ϵ	ϵ	1
prob: ϵ	1	$1 - \epsilon$	ϵ

PROPOSITION 5.1. *Assume any bidder bids her expected value when it is the dominant strategy. In Example 3, there exists a private signaling scheme that achieves arbitrarily better revenue, even in the worst Bayes Nash equilibrium, than the optimal public signaling scheme.*

Proposition 5.1 follows from Claim 5.1 and Claim 5.2. An intuitive explanation is as follows. In Example 3, bidder 1 values the item more than bidder 2 at both states. We say bidder 1 is *stronger* than bidder 2. We consider a private signaling scheme that simply reveals full information to bidder 2 and 3 but no information to bidder 1. This results in bidder 1, whose expected valuation is less than 3ϵ, to aggressively bid at least $1 - \epsilon$ in order to win bidder 2. This greatly increases

the revenue compared to the case of public signaling schemes in which case bidders bid only their expected values.

CLAIM 5.1. *The revenue obtained by the optimal public signaling scheme in Example 3 is at most 3ϵ.*

CLAIM 5.2. *In Example 3, there exists a private signaling scheme that achieves revenue at least $1 - \epsilon$ even in the worst Bayes Nash equilibrium.*

REMARK 2. *Our earlier Example 1 shows that a weak bidder may bid as low as possible if less informed, due to the winner's curse. Example 3 illustrates an interesting "dual" phenomenon: a strong bidder may bid very aggressively at equilibrium if less informed.*

Discussion of the Bidding Behavior Assumption. To obtain any non-trivial revenue guarantee in our worst case analysis, i.e., the worst Bayes Nash Equilibrium (BNE), it is *necessary* to make some bidder behavior assumptions. Otherwise, the auctioneer's revenue in the *worst* BNE of a second-price auction would always be 0 regardless of what signaling scheme she uses since that one bidder bids arbitrarily large and all other bidders bid 0 is always a BNE. To rule out these trivial equilibria, we assume throughout this section that *any bidder will bid his true value if this is the dominant strategy* (also assumed in Proposition 5.1, Claim 5.1 and 5.2). Typically, the dominant strategy equilibrium is the most ideal solution concept to adopt for revenue analysis. Unfortunately, it does not necessarily exist in second price auctions under private signaling. Our assumption can be viewed as a natural generalization of the dominant strategy equilibrium to the setting where it does not exist – i.e., bidders who have a dominant strategy will choose the dominant strategy and on top of that, we take the worst case analysis.

Proposition 5.1 shows the appealing revenue property of private signaling schemes, which motivates the computational question of designing the optimal private signaling scheme. Unfortunately, this turns out to be challenging due to several technical barriers. The first is the problem of equilibrium selection. Recall that there are usually many equilibria under private signaling schemes, as illustrated in Example 1. Since there are no natural equilibrium selection criteria, we choose to analyze the revenue in the worst-case equilibrium, as indicated in Proposition 5.1 (under the aforementioned bidding behavior assumption). Second, as illustrated by Example 1 and 3, private signaling schemes usually cause complicated bidding behaviors at equilibrium in second-price auctions. Even computing a desired equilibrium is already highly non-trivial [16, 26], let alone optimizing the revenue over these equilibria.

Nevertheless, our next result shows that, under mild assumptions, there exists a private signaling scheme that extracts almost the *full surplus* even in the *worst* Bayes Nash equilibrium. Before stating the result formally, we first specify some notations. Let $\rho^* = \max_{\theta, i} v_i(\theta)$ denote the maximum value among all possible valuations of all bidders, and i^* denotes the bidder who has value ρ^* (one of them if multiple). Let $\rho^{**} = \max_{\theta, i \neq i^*} v_i(\theta)$ be the maximum value among all possible valuations, excluding the values of bidder i^*, and i^{**} denote a bidder who has the value ρ^{**}. Let $x^+ = \max(x, 0)$. We prove the following theorem.

THEOREM 5.1. *Assume $0 \in \mathcal{V}_i$ for each i and λ supports on the* entire *set $\mathcal{V}$. There exists a private signaling scheme that can be implemented in $\mathrm{poly}(n, \sum_i |\mathcal{V}_i|)$ time and extract the full surplus, excluding an amount of $\sum_\theta \lambda_\theta [v_{i^*}(\theta) - \rho^{**}]^+ + \epsilon$ for an arbitrarily small ϵ, even in the worst Bayes Nash equilibrium.*

The intuition underlying the missed amount from the surplus in Theorem 5.1 is as follows. Recall that ρ^{**} is the largest possible value, excluding the values of bidder i^*. When i^* has a value $v_{i^*}(\theta) > \rho^{**}$ at state of nature θ, the revenue is at most ρ^{**} in a second price auction since ρ^{**} is the highest possible bid among all bidders other than i^*. Therefore, an amount of $[v_{i^*}(\theta) - \rho^{**}]$ can *not* be extracted from the surplus at such a state. The excluded amount in Theorem 5.1, up to the negligible ϵ, is precisely the sum of all such possible surplus losses weighted by their probabilities. In Example 3, $\rho^* = \rho^{**} = 1$, therefore the event $v_{i^*}(\theta) > \rho^*$ does not happen and the excluded amount equals to ϵ, which is indeed the gap between the total surplus and the revenue of our constructed private signaling scheme (Claim 5.2).

The proof of Theorem 5.1 is inspired by Example 3, which can be viewed differently from the following perspective. In particular, the second value profile $(1, 1 - \epsilon, \epsilon)$ in Example 3 can be viewed as an *auxiliary* value profile, which may have arbitrarily small probability but helps the auctioneer to extract almost full surplus from the first value profile $(2\epsilon, \epsilon, 1)$. To prove Theorem 5.1, we first generalize Example 3 by characterizing a general condition of two value profiles, under which an auxiliary value profile can be leveraged to extract (almost) full surplus from another value profile. We then show that for any realized value profile, a corresponding auxiliary value profile can always be constructed under the theorem assumptions. We defer the formal proof to Appendix D.2.

6 Conclusions and Future Work

In this paper, we study optimal signaling in second price auctions when there is a large set of possible item realizations. In a well-motivated Bayesian valuation setting, we obtain preliminary positive and negative results for the optimal public signaling. We also initiate the study of optimal private signaling in auctions and illustrate various challenges of designing the optimal private signaling scheme due to complicated equilibrium bidding behaviors and equilibrium selection issues. In the basic known-valuation setting, we exhibit a private signaling scheme that extracts almost the full surplus.

Our work has left open a rich set of interesting questions. We conclude by mentioning a few. In the Bayesian valuation setting, we proved a menu-size-complexity-type negative result. However, the computational complexity of the problem is still open. In particular, is there a polynomial-time algorithm for optimal public signaling in the Bayesian valuation setting? Our positive result here relies on the MHR assumption of bidders' valuation distributions. Is it possible to relax this assumption and obtain an approximately optimal public scheme for more general settings? Is there still a "simple" approximate scheme in the more general setting?

Our result for private signaling is still limited to the (less realistic) known valuation setting, but it illustrates the surprising power of private schemes in extracting revenue. This opens many possibilities for future research. We wonder whether similar results hold in more realistic auction settings, e.g., the Bayesian valuation setting studied in this work. In particular, to what extent can private signaling schemes outperform public schemes in terms of extracting revenue? Is the auctioneer's revenue from the optimal private scheme close to the total surplus? How to efficiently compute the optimal private signaling scheme? These descriptive and prescriptive questions are largely open in more realistic auction settings.

References

[1] Itai Arieli and Yakov Babichenko. *Private bayesian persuasion.* 2016.

[2] Moshe Babaioff, Yannai A. Gonczarowski, and Noam Nisan. *The menu-size complexity of revenue approximation.* In Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, pages 869–877, New York, NY, USA, 2017. ACM.

[3] Yakov Babichenko and Siddharth Barman. *Computational aspects of private bayesian persuasion.* 8th Innovations in Theoretical Computer Science (ITCS), 2017.

[4] Dirk Bergemann and Stephen Morris. *Bayes correlated equilibrium and the comparison of information struc-*

tures in games. Theoretical Economics, 11(2):487–522, 2016.

[5] Peter Bro Miltersen and Or Sheffet. *Send mixed signals: earn more, work less.* In Proceedings of the 13th ACM Conference on Electronic Commerce, pages 234–247. ACM, 2012.

[6] Colin M Campbell and Dan Levin. *Can the seller benefit from an insider in common-value auctions?* Journal of Economic Theory, 91(1):106–120, 2000.

[7] Yu Cheng, Ho Yee Cheung, Shaddin Dughmi, Ehsan Emamjomeh-Zadeh, Li Han, and Shang-Hua Teng. *Mixture selection, mechanism design, and signaling.* In Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on, pages 1426–1445. IEEE, 2015.

[8] Constantinos Daskalakis, Christos Papadimitriou, and Christos Tzamos. *Does information revelation improve revenue?* In Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, pages 233–250, New York, NY, USA, 2016. ACM.

[9] Shaddin Dughmi and Haifeng Xu. *Algorithmic bayesian persuasion.* In Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing, STOC '16, pages 412–425, New York, NY, USA, 2016. ACM.

[10] Shaddin Dughmi, Nicole Immorlica, and Aaron Roth. *Constrained signaling in auction design.* In Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, pages 1341–1357. Society for Industrial and Applied Mathematics, 2014.

[11] Yuval Emek, Michal Feldman, Iftah Gamzu, Renato Paes Leme, and Moshe Tennenholtz. *Signaling schemes for revenue maximization.* In Proceedings of the 13th ACM Conference on Electronic Commerce, EC '12, pages 514–531, New York, NY, USA, 2012. ACM.

[12] Hanming Fang and Sérgio O Parreiras. *Equilibrium of affiliated value second price auctions with financially constrained bidders: The two-bidder case.* Games and Economic Behavior, 39(2):215–236, 2002.

[13] Yossi Feinberg, Moshe Tennenholtz, et al. *Anonymous bidding and revenue maximization.* The BE Journal of Theoretical Economics, 5(1):1–12, 2005.

[14] Hu Fu, Patrick Jordan, Mohammad Mahdian, Uri Nadav, Inbal Talgam-Cohen, and Sergei Vassilvitskii. *Ad auctions with data.* In Algorithmic Game Theory, pages 168–179. Springer, 2012.

[15] Sergiu Hart and Noam Nisan. *The menu-size complexity of auctions.* In Proceedings of the Fourteenth ACM Conference on Electronic Commerce, EC '13, pages 565–566, New York, NY, USA, 2013. ACM.

[16] Donald B Hausch. *An asymmetric common-value auction model.* The RAND Journal of Economics, pages 611–621, 1987.

[17] Patrick Hummel and R Preston McAfee. *When does improved targeting increase revenue?* In Proceedings of the 24th International Conference on World Wide Web, pages 462–472. ACM, 2015.

[18] Emir Kamenica and Matthew Gentzkow. *Bayesian persuasion.* The American Economic Review,

101(6):2590–2615, 2011.

[19] Nitish Korula, Vahab Mirrokni, and Hamid Nazerzadeh. *Optimizing display advertising markets: Challenges and directions.* IEEE Internet Computing, 20(1):28–35, 2016.

[20] Paul Milgrom and Robert J Weber. *The value of information in a sealed-bid auction.* Journal of Mathematical Economics, 10(1):105–114, 1982.

[21] Paul R Milgrom and Robert J Weber. *A theory of auctions and competitive bidding.* Econometrica: Journal of the Econometric Society, pages 1089–1122, 1982.

[22] P.R. Milgrom. *The Structure of Information in Competitive Bidding.* Graduate School of Business, Stanford University., 1979.

[23] Motty Perry and Philip J Reny. *On the failure of the linkage principle in multi-unit auctions.* Econometrica, 67(4):895–900, 1999.

[24] Tao Qin, Wei Chen, and Tie-Yan Liu. *Sponsored search auctions: Recent advances and future directions.* ACM Transactions on Intelligent Systems and Technology (TIST), 5(4):60, 2015.

[25] Michael H Rothkopf. *A model of rational competitive bidding.* Management Science, 15(7):362–373, 1969.

[26] Vasilis Syrgkanis, David Kempe, and Eva Tardos. *Information asymmetries in common-value auctions with discrete signals.* the ACM Conference on Economics and Computation, 2015.

[27] Ina A Taneva. *Information design.* 2015.

[28] Shoshana Vasserman, Michal Feldman, and Avinatan Hassidim. *Implementing the wisdom of waze.* In Proceedings of the 24th International Conference on Artificial Intelligence, pages 660–666. AAAI Press, 2015.

[29] Zihe Wang and Pingzhong Tang. *Optimal mechanisms with simple menus.* In Proceedings of the Fifteenth ACM Conference on Economics and Computation, EC '14, pages 227–240, New York, NY, USA, 2014. ACM.

Appendix

A   Proof of Theorem 3.1

We start by exhibiting the linear program used by the algorithm, i..e, Linear Program (A.1), which uses the empirical distribution $\widetilde{\lambda}$ and relaxes the first two constraints of LP (3.1) by $\frac{\epsilon}{2n^2}$. The algorithm simply solves LP (A.1) and signals for θ as suggested by the solution of LP (A.1). Details are in Algorithm 2, which we instantiate with $\epsilon > 0$ and $K = \frac{8n^4}{\epsilon^2} \log \frac{4n^3}{\epsilon}$. It is important to notice that the signal σ^{ij} generated by Algorithm 2 does *not* necessarily correspond to the outcome where bidder i and j have the highest and second-highest values. This is because we have relaxed the constraints in LP (A.1) so that i, j are not necessarily the top two bidders. Consequently, the objective of LP (A.1) does *not* correspond to the

revenue of the auction. Nevertheless, Lemma A.1 shows that the revenue is close to the expected optimal objective value of LP (A.1), denoted as $OPT(LPA.1)$.

(A.1)
$$\max \quad \sum_{\sigma^{ij} \in \Sigma} \sum_{\theta \in \widetilde{\Theta}} \widetilde{\lambda}_\theta \varphi(\theta, \sigma^{ij}) v_j(\theta)$$
$$\text{s.t.} \quad \sum_{\theta \in \widetilde{\Theta}} \widetilde{\lambda}_\theta \varphi(\theta, \sigma^{ij}) v_i(\theta) \geq$$
$$\quad \sum_{\theta \in \widetilde{\Theta}} \widetilde{\lambda}_\theta \varphi(\theta, \sigma^{ij}) v_j(\theta) - \frac{\epsilon}{2n^2}, \quad \text{for } j \neq i.$$
$$\quad \sum_{\theta \in \widetilde{\Theta}} \widetilde{\lambda}_\theta \varphi(\theta, \sigma^{ij}) v_j(\theta) \geq$$
$$\quad \sum_{\theta \in \widetilde{\Theta}} \widetilde{\lambda}_\theta \varphi(\theta, \sigma^{ij}) v_k(\theta) - \frac{\epsilon}{2n^2}, \quad \text{for } k \neq i, j.$$
$$\quad \sum_{\sigma^{ij}} \varphi(\theta, \sigma^{ij}) = 1, \quad \quad \quad \quad \text{for } \theta.$$
$$\quad \varphi(\theta, \sigma^{ij}) \geq 0, \quad \quad \quad \quad \quad \text{for } \theta, \sigma^{ij}.$$

Figure 1: Relaxed LP using Empirical Distributions

Algorithm 2 Public Signaling Scheme for Known-Valuation Setting

---
Parameter: $\epsilon \geq 0$; Integer $K \geq 0$.
Input: State of nature θ represented by $v(\theta)$
Input: Prior distribution λ given as a sampling oracle
Output: Signal $\sigma \in \Sigma$, where $\Sigma = \left\{\sigma^{ij}\right\}_{i \neq j \in [n]}$.
1: Draw integer ℓ uniformly at random from $\{1, \ldots, K\}$, and denote $\theta_\ell = \theta$.
2: Sample $\theta_1, \ldots, \theta_{\ell-1}, \theta_{\ell+1} \ldots, \theta_K$ independently from the state of nature, and let $\widetilde{\lambda}$ denote the empirical distribution over the multiset $\widetilde{\Theta} = \{\theta_1, \ldots, \theta_K\}$.
3: Solve linear program (A.1). Let $\widetilde{\varphi} : \widetilde{\Theta} \to \Delta(\Sigma)$ be the optimal solution.
4: Output signal σ^{ij} with probability $\widetilde{\varphi}(\theta_\ell, \sigma^{ij})$.

Theorem 3.1 follows from the following two lemmas.

LEMMA A.1. *Assume $\theta \sim \lambda$, and assume bidders bid their conditional expected value upon receiving a signal. The expected revenue is at least $\mathbf{E}_{\widetilde{\Theta}} OPT(LPA.1) - \frac{\epsilon}{2}$. Both expectations are taken over the random input θ as well as internal randomness and Monte-Carlo sampling performed by the algorithm.*

Proof. Let $\mathbf{Rev}$ be the total expected revenue achieved by Algorithm 2, $\mathbf{Rev}(\sigma^{ij})$ denote the expected revenue generated by signal σ^{ij}, multiplied by the probability of receiving signal σ^{ij}. Therefore, $\mathbf{Rev} = \sum_{\sigma^{ij} \in \Sigma} \mathbf{Rev}(\sigma^{ij})$. Let $v_k(\sigma^{ij})$ denote the bidder k's value on receiving signal σ^{ij}, also multiplied by the probability of receiving signal σ^{ij}. Further, we denote $v_k(\sigma^{ij}|\widetilde{\Theta}) = \sum_{\theta \in \widetilde{\Theta}} \widetilde{\lambda}_\theta \varphi(\theta, \sigma^{ij}) v_k(\theta)$. Therefore, $v_k(\sigma^{ij}) = \mathbf{E}_{\widetilde{\Theta}} \left[v_k(\sigma^{ij}|\widetilde{\Theta})\right]$. Note that, by the princi-

ple of deferred decision, $\widetilde{\Theta} = \{\theta_1, ..., \theta_K\}$ at Step 2 can be viewed as K i.i.d. samples from λ.

The first two constraints of LP (A.1) mean $v_i(\sigma^{ij}|\widetilde{\Theta}) \geq v_j(\sigma^{ij}|\widetilde{\Theta}) - \frac{\epsilon}{2n^2}$ and $v_j(\sigma^{ij}|\widetilde{\Theta}) \geq v_k(\sigma^{ij}|\widetilde{\Theta}) - \frac{\epsilon}{2n^2}$ for all $k \neq i, j$. Taking expectation over Θ on both sides of the inequalities, we have

$$v_i(\sigma^{ij}) = \mathbf{E}_{\widetilde{\Theta}} \left[v_i(\sigma^{ij}|\widetilde{\Theta})\right]$$
(A.2)
$$\geq \mathbf{E}_{\widetilde{\Theta}} \left[v_j(\sigma^{ij}|\widetilde{\Theta}) - \frac{\epsilon}{2n^2}\right]$$
$$= v_j(\sigma^{ij}) - \frac{\epsilon}{2n^2}.$$

Similarly, $v_j(\sigma^{ij}) \geq v_k(\sigma^{ij}) - \frac{\epsilon}{2n^2}$ for all $k \neq i, j$. Since the revenue of a second-price auction is at least the minimum of any two bidders' values, we have
(A.3)
$$\mathbf{Rev}(\sigma^{ij}) \geq \min\left(v_i(\sigma^{ij}), v_j(\sigma^{ij})\right) \geq v_j(\sigma^{ij}) - \frac{\epsilon}{2n^2},$$

where the last inequality follows from Inequality (A.2). Therefore,

$$\mathbf{Rev} = \sum_{\sigma^{ij} \in \Sigma} \mathbf{Rev}(\sigma^{ij})$$
$$\geq \sum_{\sigma^{ij} \in \Sigma} \left[v_j(\sigma^{ij}) - \frac{\epsilon}{2n^2}\right]$$
$$= \sum_{\sigma^{ij} \in \Sigma} \mathbf{E}_{\widetilde{\Theta}} \left[v_j(\sigma^{ij}|\widetilde{\Theta})\right] - \frac{\epsilon}{2}$$
$$= \mathbf{E}_{\widetilde{\Theta}} \left[\sum_{\sigma^{ij} \in \Sigma} \sum_{\theta \in \widetilde{\Theta}} \widetilde{\lambda}_\theta \varphi(\theta, \sigma^{ij}) v_j(\theta)\right] - \frac{\epsilon}{2}$$
$$= \mathbf{E}_{\widetilde{\Theta}} OPT(LPA.1) - \frac{\epsilon}{2}$$

$\square$

LEMMA A.2. *Let OPT denote the expected revenue induced by the optimal signaling scheme for prior distribution λ. When Algorithm 2 is instantiated with $K = \frac{8n^4}{\epsilon^2} \log \frac{4n^3}{\epsilon}$ and its input θ is drawn from λ, the expected revenue generated by Algorithm 2 is at least $OPT - \epsilon$. Expectation is over the random input θ as well as the Monte-Carlo sampling performed by the algorithm.*

Proof. We prove that when $K = \frac{8n^4}{\epsilon^2} \log \frac{4n^3}{\epsilon}$, with probability at least $1 - \frac{\epsilon}{4}$, the expected optimal objective of LP (A.1) is at least $OPT - \frac{\epsilon}{2}$. This, together with Lemma A.1, implies that the revenue generated by Algorithm 2 is at least $OPT - \epsilon$.

Let φ^* be the solution of LP (3.1), i.e., precisely the optimal signaling scheme of the problem. At a

high level, we will show that restricting φ^* to the set $\widetilde{\Theta}$ results in a scheme that is feasible for LP (A.1) and achieves revenue at least $OPT(LPA.1) - \epsilon/2$ with high probability. Formally, let $\widetilde{\varphi}(\theta, \sigma) = \varphi^*(\theta, \sigma)$ for any $\theta \in \widetilde{\Theta}, \sigma \in \Sigma$, and $\widetilde{\lambda}$ be the empirical distribution over $\widetilde{\Theta}$. Observe that fixing any i, j, both the left-hand-side and right-hand-side terms in the first and second constraints of LP (3.1) can be viewed as the expectation of random variable $\varphi^*(\theta, \sigma^{ij})$ over random $\theta \sim \lambda$, while the corresponding constraints in LP (A.1) with constructed solution $\widetilde{\varphi}$ are precisely their empirical means over K samples. By the principle of deferred decision, these samples can be viewed as i.i.d. samples from $\theta \sim \lambda$. Therefore, with probability at least $1 - e^{-2(\frac{\epsilon}{4n^2})^2 K} > 1 - \frac{\epsilon}{4n^3}$, the empirical mean of any term is within $\frac{\epsilon}{4n^2}$ of its expectation. Since there are n different terms in total for any given σ^{ij} and $|\Sigma| = n(n-1)$, by union bound, with probability at least $1 - \epsilon/4$, all the empirical means are within $\frac{\epsilon}{4n^2}$ of their expectations. In this case, since the original expectations satisfy the first two constraints of LP (3.1), the empirical means staisfy the first (relaxed) two constraints of LP (A.1), implying that $\widetilde{\varphi}$ is feasible to LP (A.1). Moreover, the objective value of $\widetilde{\varphi}$ in LP (A.1) is within $\frac{\epsilon}{4n^2} \cdot n(n-1) \le \epsilon/4$ of the optimal objective value of LP (3.1), i.e., OPT. Therefore, we have found a feasible $\widetilde{\varphi}$ for LP (A.1) that achieves objective at least $OPT - \frac{\epsilon}{4}$ with probability at least $1 - \frac{\epsilon}{4}$. Randomness is over the samples $\widetilde{\Theta}$. This implies that $\mathbf{E}_{\widetilde{\Theta}}\big[OPT(LPA.1)\big] \ge OPT - \frac{\epsilon}{2}$, proving the lemma. $\square$

B Proof of Theorem 4.1

We construct a problem in which any scheme with $2^{o(n^{1/4})}$ signals must be far from optimality. In particular, consider a second-price auction in the Bayesian-valuation setting with n i.i.d. bidders, described as follows. Any bidder i has a valuation function V, which depends on her private type $t_i \in \{\mathbf{A}, \mathbf{B}\}$ and a binary item type $\theta_i \in \{0, 1\}$. Every bidder has the same valuation function $V(t_i, \theta_i)$, whose value table is as follows:

	$V(\cdot, 0)$	$V(\cdot, 1)$
A	0	1
B	0	0

θ_i's are i.i.d. with $\mathbf{Pr}(\theta_i = 1) = \epsilon$ for some small constant ϵ. Therefore, the prior distribution of θ is a product distribution with $\mathbf{Pr}(\theta) = \epsilon^{|\theta|}(1 - \epsilon)^{n - |\theta|}$. Moreover, bidders' private type t_i's are also i.i.d. with $\mathbf{Pr}(t_i = \mathbf{A}) = \frac{1}{\sqrt{n}}$ for each i.

For convenience, we will use D_0 to denote the point distribution at 0 and D_1 to denote the binary

distribution which takes value 1 with probability $\frac{1}{\sqrt{n}}$. Observe that in the constructed instance above, $\theta_i = 0$ or 1 specifies bidder i's valuation distribution is D_0 or D_1. We start by showing that the revenue of the optimal public signaling scheme, denoted as OPT_r, is lower bounded by a value arbitrarily close to 1 for large enough n.

LEMMA B.1. $OPT_r \ge 1 - e^{-\epsilon\sqrt{n}} - \epsilon\sqrt{n} \cdot e^{-\epsilon \frac{n-1}{\sqrt{n}}}$.

Proof. We consider the scheme of fully revealing θ to every bidder and show that the revenue of this particular public signaling scheme is already lower bounded by the RHS of the inequality in the lemma. Notice that $\mathbf{Pr}(|\theta| = i) = C_n^i \epsilon^{|\theta|}(1-\epsilon)^{n-|\theta|}$. Moreover, when $|\theta| = i$, there are i bidders who have value distribution D_1. It is not hard to check that the revenue in this case is $1 - (1 - \frac{1}{\sqrt{n}})^i - \frac{i}{\sqrt{n}}(1 - \frac{1}{\sqrt{n}})^{(i-1)}$. Therefore, the revenue of full information scheme satisfies

$$
\begin{aligned}
&\mathbf{Rev} \\
=~ &\sum_{i=0}^{n} C_n^i \epsilon^i (1-\epsilon)^{n-i} \cdot \\
&\quad \left[1 - (1 - \frac{1}{\sqrt{n}})^i - \frac{i}{\sqrt{n}}(1 - \frac{1}{\sqrt{n}})^{(i-1)}\right] \\
=~ &1 - \sum_{i=0}^{n} C_n^i \epsilon^i (1-\epsilon)^{n-i} \cdot (1 - \frac{1}{\sqrt{n}})^i - \\
&\quad \sum_{i=0}^{n} C_n^i \epsilon^i (1-\epsilon)^{n-i} \cdot \frac{i}{\sqrt{n}}(1 - \frac{1}{\sqrt{n}})^{(i-1)} \\
=~ &1 - \left(1 - \epsilon + \epsilon - \frac{\epsilon}{\sqrt{n}}\right)^n - \\
&\quad \epsilon\sqrt{n} \sum_{i=0}^{n} C_{n-1}^{i-1}(1-\epsilon)^{n-i} \cdot \epsilon^{i-1}(1 - \frac{1}{\sqrt{n}})^{(i-1)} \\
=~ &1 - \left(1 - \frac{\epsilon}{\sqrt{n}}\right)^n - \epsilon\sqrt{n}\left(1 - \epsilon + \epsilon - \frac{\epsilon}{\sqrt{n}}\right)^{n-1} \\
\ge~ &1 - e^{-\epsilon\sqrt{n}} - \epsilon\sqrt{n} \cdot e^{-\frac{\epsilon(n-1)}{\sqrt{n}}}.
\end{aligned}
$$

Therefore, $OPT_r \ge \mathbf{Rev} \ge 1 - e^{-\epsilon\sqrt{n}} - \epsilon\sqrt{n} \cdot e^{-\frac{\epsilon(n-1)}{\sqrt{n}}}$, as desired. $\square$

We will now upper bound the revenue of any signaling scheme that uses N signals, and show that the revenue is small when $N = 2^{o(n^{1/4})}$. Unfortunately, directly examining revenue is difficult. We instead consider an obvious upper bound of revenue, i.e., the welfare, and instead upper bound the maximum welfare of any signaling scheme using N signals.

We say a signaling scheme is *deterministic* if it maps each state of nature deterministically to one signal. As a result, each signal corresponds to a subset of

states and all such subsets form a partition of the space of states of nature. A deterministic signaling scheme with N signals is denoted by $\{S_1, ..., S_N\}$, which is a partition of $\{0,1\}^n$. We therefore also call S_i a signal. The following lemma shows that there always exists a deterministic signaling scheme that achieves the maximum welfare. Intuitively, this is because welfare is a convex function (more precisely, the max function) of bidders' valuations, thus achieves its optimal value at a vertex of the polytope of signaling schemes.

LEMMA B.2. *[10] Let $\mathcal{X}_N$ denote all possible public signaling schemes using at most N signals. There always exists a deterministic signaling scheme that achieves the maximum welfare among schemes in $\mathcal{X}_N$. Such a deterministic signaling scheme partitions the set $\{0,1\}^n$ into at most N subsets.*

As a result of Lemma B.2, it is without loss of generality to upper bound the welfare achieved by deterministic signaling schemes, which are collections of subsets of $\{0,1\}^n$. For any $S \subseteq \{0,1\}^n$, we use $\mathbf{Pr}(S) = \sum_{\theta \in S} \lambda_\theta$ to denote the probability mass of the subset. At a high level, our idea is to show that any subset of $\{0,1\}^n$ either results in a welfare of $\mathcal{O}(\epsilon)$ or has exponentially small probability mass (specifically, $\mathbf{Pr}(S) \leq 2^{-\Theta(n^{1/4})}$). This will imply that a deterministic scheme with $2^{o(n^{1/4})}$ signals (i.e., subsets) must have $\mathcal{O}(\epsilon)$ welfare.

For any $S \subseteq \{0,1\}^n$, let

$$\overline{\theta}(S) = \mathbf{E}(\theta | \theta \in S) = \frac{\sum_{\theta \in S} \epsilon^{|\theta|}(1-\epsilon)^{n-|\theta|} \cdot \theta}{\sum_{\theta \in S} \epsilon^{|\theta|}(1-\epsilon)^{n-|\theta|}}$$

denote the expectation of the states in S. When S is clear from the context, we simply use $\overline{\theta}$ to denote the expected state. Observe that $\overline{\theta}_i$ is the probability that bidder i's valuation function is D_1 conditioned on receiving the signal S. The following lemma shows that if $\overline{\theta}(S)$ have many large components, then $\mathbf{Pr}(S)$ should be small.

LEMMA B.3. *For any $S \subseteq \{0,1\}^n$, let m be the number of entries in $\overline{\theta}(S)$ with values at least 2ϵ. Then we have $\mathbf{Pr}(S) \leq e^{-2\sqrt{m}}$.*

Proof. By symmetry, w.l.o.g., we assume the first m entries of $\overline{\theta}$ are at least 2ϵ. We view this as a constrain on the set S and will seek to maximize $\mathbf{Pr}(S)$ over all possible sets S subject to the constraint. This can be abstractly formulated as an optimization problem.

(B.4)
$$\begin{aligned} \max \quad & \mathbf{Pr}(S) = \sum_{\theta \in S} \epsilon^{|\theta|}(1-\epsilon)^{n-|\theta|} \\ \text{s.t.} \quad & S : \text{values of first } m \text{ entries of } \overline{\theta}(S) \geq 2\epsilon. \end{aligned}$$

Program for Maximizing $\mathbf{Pr}(S)$

Unfortunately, Optimization Program (B.4) is difficult to reason about. We instead relax the constraint to be that the first m entries sum up to at least $2m\epsilon$ and the consider the relaxed program as follows.

(B.5)
$$\begin{aligned} \max \quad & \mathbf{Pr}(S) = \sum_{\theta \in S} \epsilon^{|\theta|}(1-\epsilon)^{n-|\theta|} \\ \text{s.t.} \quad & \text{Sum of first } m \text{ entries of } \overline{\theta} \geq 2m\epsilon. \end{aligned}$$

Relaxed Program for Maximizing $\mathbf{Pr}(S)$.

It turns out that the optimal solution to Optimization Program (B.5) can be explicitly constructed. In particular, the program seeks to maximize $\mathbf{Pr}(S)$ subject to that the sum of first m entries have large expected values. Some thoughts reveal that the optimal solution are obtained as follows: add all the θ with first m entries summing up to m, then add all the θ with first m entries summing up to $m-1$, and so on so forth, until that the expected sum of the first m entries is less than $2m\epsilon$. Let k denote the sum of the first m entries of the last element added to S. Obviously, $k < 2m\epsilon$. We now argue that k is at least $\epsilon m + m^{3/4}$. This requires a technical lemma which will be proved at the end of this section.

LEMMA B.4. *Let $X_1, X_2, ..., X_m$ be m i.i.d. Bernoulli random variables with $\mathbf{E}(X_i) = \epsilon$ for every i and $X = \sum_{i=1}^{m} X_i$. Let $k = \epsilon m + m^{3/4}$ and $\epsilon \in (0, \frac{1}{2})$, there exists N such that $\mathbf{E}(X | X \geq k) = \frac{\sum_{i=k}^{m} C_m^i \epsilon^i (1-\epsilon)^{m-i} \cdot i}{\sum_{i=k}^{m} C_m^i \epsilon^i (1-\epsilon)^{m-i}} < 2m\epsilon$ for any $m \geq N$. Here $C_m^i = \frac{m!}{i!(m-i)!}$ is the binomial coefficient.*

Lemma B.4 shows that k is at least $\epsilon m + m^{3/4}$, since otherwise the expected sum of the first m entries of $\overline{\theta}(S)$ is less than $2m\epsilon$, violating the constraint of Program B.5. As a result, $OPT(PB.5)$ is upper bounded by $P(k) = \sum_{i=k}^{m} C_m^i \epsilon^i (1-\epsilon)^{m-i}$, which is the probability that sum of m i.i.d. binary random variables is at least k. By the Hoeffding bound, we know that $P(k) \leq e^{-2\sqrt{m}}$ when $k \geq \epsilon m + m^{3/4}$. Therefore, $OPT(PB.4) \leq OPT(B.5) \leq P(\epsilon m + m^{3/4}) \leq e^{-2\sqrt{m}}$, concluding the proof. $\square$

Utilizing Lemma B.3, we can now formally show that for any signal $S \subseteq \{0,1\}^n$, either the welfare from S is $\mathcal{O}(\epsilon)$ or its probability is exponentially small.

LEMMA B.5. *For any $S \subseteq \{0,1\}^n$, at least one of the following two holds:*

1. *The conditional welfare of signal S is at most 3ϵ.*

2. $\mathbf{Pr}(S) \leq e^{-2n^{1/4}C}$ *where* $C = \sqrt{\ln \frac{1-2\epsilon}{1-3\epsilon}} > 0$ *is a constant depending on ϵ.*

Proof. Given $\bar{\theta}$, the (expected) value distribution for bidder i is $\bar{\theta}_i$ with probability $1/\sqrt{n}$ and 0 otherwise. Let m denote the number of entries of $\bar{\theta}$ that are at least 2ϵ. By independence, the probability that at least one bidder's value is at least 2ϵ is $1 - (1 - \frac{1}{\sqrt{n}})^m$. Therefore, the conditional welfare $\mathbf{Wel}(S)$ is upper bounded $2\epsilon \cdot (1 - \frac{1}{\sqrt{n}})^m + 1 - (1 - \frac{1}{\sqrt{n}})^m$. We show that if $\mathbf{Wel}(S) > 3\epsilon$, then $m > \sqrt{n} \cdot \ln \frac{1-2\epsilon}{1-3\epsilon}$, as follows

$$\mathbf{Wel}(S) > 3\epsilon$$

$$\Rightarrow \quad 2\epsilon \cdot (1 - \frac{1}{\sqrt{n}})^m + 1 - (1 - \frac{1}{\sqrt{n}})^m > 3\epsilon$$

$$\Rightarrow \quad 1 - 3\epsilon > (1 - 2\epsilon) \cdot (1 - \frac{1}{\sqrt{n}})^m$$

$$\Rightarrow \quad \ln \frac{1-3\epsilon}{1-2\epsilon} > m \ln(1 - \frac{1}{\sqrt{n}}) \geq m \cdot (-\frac{1}{\sqrt{n}})$$

$$\Rightarrow \quad m > \sqrt{n} \cdot \ln \frac{1-2\epsilon}{1-3\epsilon}$$

Therefore, if $\mathbf{Wel}(S) > 3\epsilon$, we have by Lemma B.3 that

$$\mathbf{Pr}(S) \leq e^{-2\sqrt{m}} \leq e^{-2n^{1/4}C}$$

for $C = \sqrt{\ln \frac{1-2\epsilon}{1-3\epsilon}}$. $\qquad\square$

Finally, let $\mathcal{S} = \{S_1, ..., S_N\}$ be any deterministic signaling scheme with N signals. Utilizing Lemma B.5, we have

$$\mathbf{Wel}(\mathcal{S})$$
$$= \sum_{S : \mathbf{Wel}(S) \leq 3\epsilon} \mathbf{Wel}(S) \mathbf{Pr}(S) + \sum_{S : \mathbf{Wel}(S) > 3\epsilon} \mathbf{Wel}(S) \mathbf{Pr}(S)$$
$$\leq \sum_{S : \mathbf{Wel}(S) \leq 3\epsilon} 3\epsilon \cdot \mathbf{Pr}(S) + \sum_{S : \mathbf{Wel}(S) > 3\epsilon} \mathbf{Wel}(S) \cdot e^{-2n^{1/4}C}$$
$$\leq 3\epsilon + N \cdot e^{-2n^{1/4}C},$$

where constant $C = \sqrt{\ln \frac{1-2\epsilon}{1-3\epsilon}}$. For any $N = 2^{o(n^{1/4})}$, $3\epsilon + N \cdot e^{-2n^{1/4}C}$ is upper bounded by 4ϵ for large enough n. On the other hand, the revenue is close to 1. Since ϵ can be an arbitrarily small constant, any signaling scheme with $2^{o(n^{1/4})}$ signals cannot guarantee a constant fraction of the optimal revenue for any constant less than 1. This concludes our proof of Theorem 4.1.

Proof of the Technical Lemma B.4.
(Note: we use p instead of ϵ as in the original lemma statement.)
Lemma Statement: Let $X_1, X_2, ..., X_m$ be m i.i.d. Bernoulli random variables with $\mathbf{E}(X_i) = p$ for every i and $X = \sum_{i=1}^m X_i$. Let $k = pm + m^{3/4}$ and $p \in (0, \frac{1}{2})$, there exists N such that $\mathbf{E}(X | X \geq k) = \frac{\sum_{i=k}^m C_m^i p^i (1-p)^{m-i} \cdot i}{\sum_{i=k}^m C_m^i p^i (1-p)^{m-i}} < 2mp$ for any $m \geq N$, where C_m^i is the binomial coefficient.

For convenience, we define the following terms:

$$C(m, k; p) = \mathbf{Pr}(X = k) = C_m^k p^k (1-p)^{m-k}.$$

$$S(m, k; p) = \mathbf{Pr}(X \geq k) = \sum_{i=k}^m C_m^i p^i (1-p)^{m-i}$$

$$E(m, k; p) = \mathbf{E}(X | X \geq k) \cdot \mathbf{Pr}(X \geq k)$$
$$= \sum_{i=k}^m C_m^i p^i (1-p)^{m-i} \cdot i$$

LEMMA B.6. *Fix p, there exists large enough N such that for any integer $m \geq N$ and $k \in [pm, pm + m^{3/4} + 1]$, we have*

$$S(m, k+1; p) > C(m, k; p).$$

Proof. We first upper bound the ratio between $C(m, k+1; p)$ and $C(m, k; p)$ for $k \in [pm, pm + m^{3/4} + 3]$.

$$r_k = \frac{C(m, k+1; p)}{C(m, k; p)}$$
$$= \frac{C_m^{k+1} p^{k+1} (1-p)^{m-k-1}}{C_m^k p^k (1-p)^{m-k}}$$
$$= \frac{m-k}{k+1} \cdot \frac{p}{1-p}$$
$$\geq \frac{m - pm - m^{3/4} - 3}{pm + m^{3/4} + 4} \cdot \frac{p}{1-p}$$

Note that $\frac{m-pm-m^{3/4}-3}{pm+m^{3/4}+4} \leq \frac{m-pm}{pm} = \frac{1-p}{p}$, and tends to $\frac{1-p}{p}$ as $m \to \infty$. Therefore, there exists large enough N such that $\frac{n-pm-m^{3/4}-3}{pm+m^{3/4}+4} \geq \frac{3}{4} \cdot \frac{1-p}{p}$, thus $r_k \geq \frac{3}{4}$, for any $m \geq N$. Therefore, for any $k \in [pm, pm + m^{3/4} + 1]$.

$$S(m, k+1; p) > C(m, k+1; p) + C(m, k+2; p)$$
$$= r_k C(m, k; p) + r_k r_{k+1} C(m, k; p)$$
$$\geq \left[\frac{3}{4} + (\frac{3}{4})^2\right] \cdot C(m, k; p)$$
$$\geq C(m, k; p).$$

This shows that $S(m, k+1; p) > C(m, k; p)$ for any $m \geq N$, as desired. $\qquad\square$

Proof of the Lemma:

Note that $\mathbf{E}(X|X \geq k) = \frac{E(m,k;p)}{S(m,k;p)}$. Let $\Delta(m,k;p) = E(m,k;p) - 2pmS(m,k;p)$. We prove that $\Delta(m,k;p) < 0$.

$$\frac{\Delta(m,k;p)}{m}$$

$$= \sum_{i=k}^{m} \frac{i}{m} \cdot C_m^i p^i (1-p)^{m-i} - 2p \sum_{i=k}^{m} C_m^i p^i (1-p)^{m-i}$$

$$= \sum_{i=k}^{m} C_{m-1}^{i-1} p^i (1-p)^{m-i}$$

$$\quad - 2p \cdot \sum_{i=k}^{m} \left[C_{m-1}^{i-1} + C_{m-1}^i \right] p^i (1-p)^{m-i}$$

$$= (1-2p) \cdot \sum_{i=k}^{m} C_{m-1}^{i-1} p^i (1-p)^{m-i}$$

$$\quad - 2p \cdot \sum_{i=k}^{m} C_{m-1}^i p^i (1-p)^{m-i}$$

$$= (1-2p)p \cdot \left[S(m-1,k;p) + C(m-1,k-1;p) \right]$$

$$\quad - 2p(1-p) \cdot S(m-1,k;p)$$

where the last step of derivation omitted some calculations which follows basic algebraic manipulations. Note that $m^{3/4} < (m-1)^{3/4} + 1$ for any $m > 1$, therefor $k-1 = pm - 1 + m^{3/4} < p(m-1) + (m-1)^{3/4} + 1$. It is easy to see that $k - 1 > p(m-1)$. As a result, $k - 1 \in [p(m-1), p(m-1) + (m-1)^{3/4} + 1]$. By Lemma B.6, there exists N such that $S(m-1,k;p) > C(m-1,k-1;p)$ for any $m \geq N$. In this case, we have

$$\frac{\Delta(m,k;p)}{m} < S(m-1,k;p) \left[(1-2p)2p - 2p(1-p) \right] \leq 0.$$

Therefore, $\Delta(m,k:p) < 0$ for any $m \geq N$.

C Omitted Proofs from Section 4

C.1 Proof of Lemma 4.1. Since any $\theta \in \Theta_1$ can be viewed as a bidder i, we will think of Θ_1 as $[n]$ and a bidder i as a state in $[n]$. Let p_i denote the probability of state i. W.l.o.g., assume that $p_1 \geq p_2 ... \geq p_n$. We now prove each direction.

" $\Rightarrow$ ": Since $p_1 \cdot \pi(1,i) = p_i \cdot \pi(i,1)$ for any $i = 1, ... n$ (note $\pi(1,1) = 0$), summing over i, we obtain

$$p_1 = \sum_{i=1}^{n} p_1 \cdot \pi(1,i) = \sum_{i=1}^{n} p_i \cdot \pi(i,1)$$

$$= \sum_{i=2}^{n} p_i \cdot \pi(i,1) \leq \sum_{i=2}^{n} p_i.$$

Since $p_1 \geq p_j$ for any j, we have $p_j \leq \sum_{i \neq j} p_i$, as desired.

" $\Leftarrow$ ": We show that a pooling scheme can be explicitly constructed. The pairing process can be thought of matching probability mass between pairs of bidders. First observe that if $p_1 = p_2$, there always exists a feasible pooling scheme. This is because we can match bidder n with bidder $n-1$ for probability mass of p_n so that the left probability mass of bidder $n-1$ is $p'_{n-1} = p_{n-1} - p_n$ and bidder n has 0 probability mass left. We then match bidder $n-1$ with bidder $n-2$ for probability mass p'_{n-1}. This process can be continued until only bidder $1, 2, 3$ have non-zero probability masses, satisfying $p_1 = p_2 \geq p'_3$. Now we can match bidder 1 with bidder 3 for probability mass $p'_3/2$, bidder 2 with bidder 3 for probability mass $p'_3/2$, and bidder 1 with bidder 2 for probability mass $p_1 - p'_3/2$, yielding a feasible pooling scheme.

When $p_1 > p_2$, we first gradually match bidder 1 with bidder 2 until the left probability masses of bidder 2 and 3 are equal. Then we gradually match bidder 1 with bidder 2 and bidder 1 with bidder 3 simultaneously at the same rate of probability mass, until one of the following happens: either bidder 1 has the same probability mass as bidder 2 and 3 or bidder 2 and 3 have the same probability mass as bidder 4. In the first case, a feasible pooling scheme exists for the left probability masses because the probability mass of bidder 1 and 2 now equal. In the second case, we continue to gradually match bidder 1 with bidder $2, 3, 4$ simultaneously at the same rate. Since $p_1 \leq \sum_{i=2}^{n} p_i$, this process will stop at some point with bidder 1 and 2 having the same probability mass, in which case a pooling scheme can then be constructed using the earlier procedure.

C.2 Proof of Theorem 4.2 and Lemma 4.2.
Recall that the welfare is maximized by revealing full information, and the optimal welfare is the expectation of the welfare at each state of nature. Our proof will lower bound the revenue in terms of the optimal welfare. In fact, we will prove an even stronger result: given any realized state of nature, the revenue generated at that state is at least $1/8$ fraction of the welfare at that state. We prove separately for the states that are fully revealed and these that are pooled.

Our analysis involves various comparisons of welfare and revenue in the standard second price auction (with no signaling). The main technical barrier comes from the fact that bidders are *asymmetric* in the sense that they have either the high or low distribution at any given state of nature. This makes it difficult to analytically characterize revenue and welfare. Our strategy is

to first analyze the *symmetric* case in which all bidders have the same value distribution $H(v)$, and then extend the bound to asymmetric bidders. We use A_k to denote a standard second price auction with k i.i.d. bidders, and $\mathbf{Rev}(A_k; H)$ and $\mathbf{Wel}(A_k; H)$ denote the revenue and welfare respectively when bidders in A_k have value distribution $H(v)$. Let $A_{n,k}$ denote a standard second price auction with n independent bidders, among whom k bidders have value high distribution $H(v)$ and another $n - k$ bidders have low value distribution $L(v)$. $\mathbf{Rev}(A_{n,k})$ and $\mathbf{Wel}(A_{n,k})$ denote the revenue and welfare of the auction $A_{n,k}$, respectively. The notation $v_{[1]}$ and $v_{[2]}$ are reserved for the largest and second largest bidder value in the auction.

In auction A_k with distribution $H(v)$, $\mathbf{Pr}(v_{[1]} \leq v) = H^k(v)$ and $\mathbf{Pr}(v_{[2]} \leq v) = kH^{k-1}(v) - (k-1)H^k(v)$. The welfare and revenue is the expectation of $v_{[1]}$ and $v_{[2]}$ respectively, which can be expressed as follows. Here, we use constant $C_k = 1 + \frac{1}{2} + \ldots \frac{1}{k}$ denote the harmonic number.

$$
\begin{aligned}
&\mathbf{Wel}(A_k; H) \\
=~ &\int_0^\infty [1 - H^k(v)]dv \\
=~ &\int_0^\infty \frac{1 - H(v)}{h(v)}\left[1 + H(v)\ldots + H^{k-1}(v)\right]h(v)dv \\
=~ &\frac{1 - H(v)]}{h(v)}\left[H(v) + \frac{H^2(v)}{2}\ldots + \frac{H^k(v)}{k}\right]\Big|_0^\infty - \\
&\int_0^\infty \left(\frac{1 - H(v)]}{h(v)}\right)'\left[H(v) + \frac{H^2(v)}{2}\ldots + \frac{H^k(v)}{k}\right]dv \\
=~ &\lim_{v\to\infty} \frac{1 - H(v)}{h(v)} \cdot C_k + \\
&\int_0^\infty \left(-\frac{1 - H(v)]}{h(v)}\right)'\left[H(v) + \frac{H^2(v)}{2}\ldots + \frac{H^k(v)}{k}\right]dv
\end{aligned}
$$

and similarly

$$
\begin{aligned}
&\mathbf{Rev}(A_k; H) \\
=~ &\int_0^\infty [1 - kH^{k-1}(v) + (k-1)H^k(v)]dv \\
=~ &\int_0^\infty \frac{1 - H}{h}\left[1 + H\ldots + H^{k-2} - (k-1)H^{k-1}\right]hdv \\
=~ &\lim_{v\to\infty} \frac{1 - H(v)}{h(v)} \cdot (C_k - 1) + \int_0^\infty \left(-\frac{1 - H(v)]}{h(v)}\right)' \cdot \\
&\left[H(v) + \frac{H^2(v)}{2}\ldots + \frac{H^{k-1}(v)}{k-1} - \frac{(k-1)H^k(v)}{k}\right]dv
\end{aligned}
$$

The last equality in both expressions comes from integration by parts. For convenience, let $x = H(v) \in [0, 1]$. The term in the square bracket of $\mathbf{Rev}(A_k; H)$ is

denoted as $R_k(x) = x + x/2 + \ldots, x^k/k - x^k$, while the term in the square bracket of $\mathbf{Wel}(A_k; H)$ is denoted as $W_k(x) = x + x/2 + \ldots, x^k/k$. The following lemma is central to our analysis.

LEMMA C.1. *For any integer $k, i > 0$, if there exists $\rho > 0$ such that $C_k - 1 \geq \rho \cdot C_i$ and $R_k(x) \geq \rho \cdot W_i(x)$ for all $x \in [0, 1]$, then $\mathbf{Rev}(A_k; H) \geq \rho \cdot \mathbf{Wel}(A_i; H)$.*

Proof. Notice that $\frac{1 - H(v)}{h(v)} \geq 0$ by definition, and $(-\frac{1 - H(v)]}{h(v)})' \geq 0$ by the MHR assumption. Comparing the terms in $\mathbf{Rev}(A_k; H)$ and $\mathbf{Wel}(A_i; H)$, it is not hard to conclude that the assumed conditions are sufficient to guarantee $\mathbf{Rev}(A_k; H) \geq \rho \cdot \mathbf{Wel}(A_i; H)$. $\square$

Using Lemma C.1, we now prove that when $|\theta|_1 \geq 2$ or $|\theta|_1 = 0$, in which case we fully reveal θ and then run a standard second price auction, the revenue is at least $\frac{1}{8}$ fraction of welfare.

LEMMA C.2. $\mathbf{Rev}(A_k; H) \geq \max\{\frac{1}{3}, \frac{\ln(k+1)-1}{\ln(k+1)}\} \cdot \mathbf{Wel}(A_k; H)$ *for any $k \geq 2$.*

Proof. It is easy to check

$$
f(x) = \frac{R_k(x)}{W_k(x)} = 1 - \frac{x^k}{x + x^2/2 + \ldots x^k/k}
$$

is a decreasing function for $x \in [0, 1]$, therefore, we have

$$
f(x) \geq f(1) = \frac{C_k - 1}{C_k}
$$

That is, $R_k(x) \geq \frac{C_k-1}{C_k}W_k(x)$. Invoking Lemma C.1, we have

$$
\frac{\mathbf{Rev}(A_k; H)}{\mathbf{Wel}(A_k; H)} \geq \frac{C_k - 1}{C_k} \geq \frac{1}{3}
$$

where we used the inequality $C_k = 1 + \frac{1}{2}\ldots + \frac{1}{k} \geq 3/2$ when $k \geq 2$ and $\frac{x-1}{x}$ is increasing in x. Similarly, $\mathbf{Rev}(A_k; H) \geq \frac{\ln(k+1)-1}{\ln(k+1)} \mathbf{Wel}(A_k; H)$ since $C_k \geq \ln(k+1)$. This proves the lemma. $\square$

Lemma C.2 implies that when $|\theta| = 0$, in which case we reveal full information, the revenue is at least $\max\{\frac{1}{3}, \frac{\ln(k+1)-1}{\ln(k+1)}\}$ fraction of the welfare (Part 1 of Lemma 4.2). We now prove similar guarantee for the case $|\theta| \geq 2$. This is proven by extending the above bound to the case with asymmetric bidders.

LEMMA C.3. *[Part 2 of Lemma 4.2] For $k \geq 2$, $\mathbf{Rev}(A_{n,k}) \geq \max\{\frac{1}{6}, \frac{\ln(k+1)-1}{2\ln(k+1)}\} \cdot \mathbf{Wel}(A_{n,k})$.*

Proof. Let $m = n - k$ denote the number of low-distribution bidders, $\mathbf{u} \in \mathcal{R}^k, \mathbf{v} \in \mathcal{R}^m$ denote

the value profile of the high-distribution bidders and low-distribution bidders respectively. Notice that $\max(\mathbf{u}, \mathbf{v}) \leq \max(\mathbf{u}) + \max(\mathbf{v})$ for any $\mathbf{u} \in \mathcal{R}^k, \mathbf{v} \in \mathcal{R}^m$, we therefore have

$$
\begin{aligned}
\mathbf{Wel}(A_{n,k}) &= \underset{\mathbf{u} \sim H^k}{\mathbf{E}} \underset{\mathbf{v} \sim L^m}{\mathbf{E}} \max(\mathbf{u}, \mathbf{v}) \\
&\leq \underset{\mathbf{u} \sim H^k}{\mathbf{E}} \underset{\mathbf{v} \sim L^m}{\mathbf{E}} [\max(\mathbf{u}) + \max(\mathbf{v})] \\
&= \mathbf{Wel}(A_k; H) + \mathbf{Wel}(A_m; L)
\end{aligned}
$$

We now lower bound $\mathbf{Rev}(A_{n,k})$ in terms of $\mathbf{Wel}(A_{n,k})$. Since the auction consists of k high-distribution bidders and m low-distribution bidders, we have $\mathbf{Rev}(A_{n,k}) \geq \mathbf{Rev}(A_k; H)$ and $\mathbf{Rev}(A_{n,k}) \geq \mathbf{Rev}(A_m; L)$. When $k \leq m$, we have

$$
\begin{aligned}
\mathbf{Rev}(A_{n,k}) &\geq \frac{1}{2}\big(\mathbf{Rev}(A_m; L) + \mathbf{Rev}(A_k; H)\big) \\
&\geq \frac{1}{2}[\max\{\frac{1}{3}, \frac{\ln(m+1)-1}{\ln(m+1)}\} \cdot \mathbf{Wel}(A_m; L) \\
&\quad + \max\{\frac{1}{3}, \frac{\ln(k+1)-1}{\ln(k+1)}\} \cdot \mathbf{Wel}(A_k; H)] \\
&\geq \max\{\frac{1}{6}, \frac{\ln(k+1)-1}{2\ln(k+1)}\}\mathbf{Wel}(A_{n,k}).
\end{aligned}
$$

where the second inequality follows from Lemma C.2. When $k \geq m$, we similarly have

$$
\begin{aligned}
\mathbf{Rev}(A_{n,k}) &\geq \mathbf{Rev}(A_k; H) \\
&\geq \max\{\frac{1}{3}, \frac{\ln(k+1)-1}{\ln(k+1)}\}\mathbf{Wel}(A_k; H) \\
&\geq \max\{\frac{1}{6}, \frac{\ln(k+1)-1}{2\ln(k+1)}\} \mathbf{Wel}(A_{n,k}).
\end{aligned}
$$

$\square$

So far we have shown that at the states that are fully revealed, the revenue is at least $\frac{1}{6}$ fraction of the welfare. We now prove similar guarantees for the states that are pooled. Note that when Algorithm 1 outputs a pool signal $(\theta', pool)$ where $\theta' \in \Theta_2$, the two bidders $i, j \in \theta'$ are equally likely (by definition of the 1-tail balanced pooling scheme), therefore each bidder in θ' has a high distribution with probability $1/2$. Assume bidders bid in expectation, the revenue in this case is at least $\frac{1}{2} \cdot \mathbf{Rev}(A_{n,2})$. The following lemma shows that $\frac{1}{2}\mathbf{Rev}(A_{n,2}) \geq \frac{1}{8}\mathbf{Wel}(A_{n,1})$. This implies that in Algorithm 1, the revenue at any pooled state is at least $\frac{1}{8}$ fraction of the welfare of that state.

LEMMA C.4. *[Part 3 of Lemma 4.2] When $n \geq 12$, we have $\frac{1}{2} \cdot \mathbf{Rev}(A_{n,2}) \geq \frac{1}{8} \cdot \mathbf{Wel}(A_{n,1})$.*

Proof. We first prove $\mathbf{Rev}(A_2; H) \geq \frac{1}{2}\mathbf{Wel}(A_1; H)$. This is because $C_2 - 1 = \frac{1}{2} = \frac{1}{2}C_1$ and $\frac{R_2(x)}{W_1(x)} = \frac{x - \frac{x^2}{2}}{x} \geq \frac{1}{2}$ for any $x \in [0,1]$. Invoking Lemma C.1, we obtain the conclusion.

Now we argue that when $n \geq 12$, $\mathbf{Rev}(A_{n-2}; H) \geq \frac{1}{2}\mathbf{Wel}(A_{n-1}; H)$. Note that $C_{n-1} \leq 2[C_{n-2} - 1]$ for any $n \geq 12$. By Lemma C.1, we only need to prove $W_{n-1}(x) \leq 2R_{n-2}(x)$ for any $x \in [0,1]$ and $n \geq 12$, as follows:

$$
\begin{aligned}
\frac{R_{n-2}(x)}{W_{n-1}(x)} &= \frac{x + x/2 + ..., x^{n-2}/(n-2) - x^{n-2}}{x + x/2 + ..., x^{n-1}/(n-1)} \\
&= 1 - \frac{x^{n-1}/(n-1) + x^{n-2}}{x + x/2 + ..., x^{n-1}/(n-1)} \\
&\geq 1 - \frac{(\frac{1}{n-1} + 1)x^{n-2}}{x + x/2 + ..., x^{n-2}/(n-2)} \\
&\geq \frac{C_{n-2} - 1 - \frac{1}{n-1}}{C_{n-2}} \\
&\geq \frac{1}{2}
\end{aligned}
$$

Similar to the proof in Lemma C.3, we have $\mathbf{Wel}(A_{n,1}) \leq \mathbf{Wel}(A_1; H) + \mathbf{Wel}(A_{n-1}; L)$ and $\mathbf{Rev}(A_{n,2}) \geq \frac{1}{2}(\mathbf{Rev}(A_2; H) + \mathbf{Rev}(A_{n-2}; L))$. Therefore

$$
\begin{aligned}
\mathbf{Rev}(A_{n,2}) &\geq \frac{1}{2} \cdot [\frac{1}{2}\mathbf{Wel}(A_1; H) + \frac{1}{2}\mathbf{Wel}(A_{n-1}; L)] \\
&\geq \frac{1}{4}\mathbf{Wel}(A_{n,1}).
\end{aligned}
$$

$\square$

D Omitted Proofs from Section 5

D.1 Proof of Claim 5.1 and Claim 5.2.

Proof of Claim 5.1. Note that in public signaling schemes, bidders always bid their expected values by our behavior assumption. In Example 3, bidder 1's valuation is higher than bidder 2's at any state, therefore bidder 1's bid is always greater than bidder 2's, conditioned on any signal. This implies that bidder 1's bid is either the highest or the second highest given any signal. As a result, the revenue is at most bidder 1's bid, which in expectation is less than 3ϵ.

Proof of Claim 5.2. Consider the following private signaling scheme: the auctioneer informs bidder 2 and 3 the true state but reveals no information to bidder 1. Since bidder 2 and 3 know the exact state, bidding the true value is their dominant strategy. We claim that the equilibrium bid of bidder 1 lies in the interval $[1-\epsilon, 1)$. To see this, observe that any bid less than $1-\epsilon$ does not place bidder 1 in the winning position in neither state,

thus results in expected utility 0 to bidder 1. However, any bid $b \in (1 - \epsilon, 1)$ helps bidder 1 to win the auction at the second state, which results in a strictly positive utility for bidder 1. Therefore, bidder 1 bids at least $1 - \epsilon$ in any equilibrium, and the auctioneer's revenue is also at least $1 - \epsilon$, concluding the proof.

D.2 Proof of Theorem 5.1. Example 3 illustrates a situation in which the auctioneer can use a private signaling scheme to extract revenue of at least $1 - \epsilon$, which is close to the total surplus 1. Interestingly, this example can be thought differently from the perspective of signaling schemes. In particular, imagine we are given any realized value profile $(2\epsilon, \epsilon, 1)^T$, Example 3 and Claim 5.2 show that by "mixing" this value profile with another value profile $(1, 1 - \epsilon, \epsilon)^T$ of arbitrarily small probability, we can extract almost full surplus from the given value profile $(2\epsilon, \epsilon, 1)^T$ using private signaling schemes. If we can do such mixing for any realized value profile, the aggregated revenue should be close to the full surplus. It turns out that this insight can be made precise.

We start by generalizing Example 3 and characterizing the structure of value profiles that leads to high revenue. Throughout the proof, we will use u_i to denote the i'th entry of any vector u and $u_{[i]}$ to denote the i'th largest entry of u. The following lemma plays a key role in our poof.

LEMMA D.1. *Consider a second-price auction in the known-valuation setting with two possible value profiles $v, u \in \mathbb{R}^n$ that have the following structure:*

	Probability	
	v : prob. $(1-\delta)$	u : prob. δ
...	...	...
bidder j	...	$u_{[1]} \ (> u_{[2]})$
...	...	...
bidder k	...	$u_{[2]} \ (= v_{[1]} - \epsilon)$
...	...	...
bidder i	$v_{[1]}$	...
...	...	...

If $j \neq i$ (but k could equal i) and $v_{[1]}, u_{[1]}, u_{[2]}$ satisfy the conditions in the above table , then for any $\delta > 0$, the following private signaling scheme extracts revenue at least $(v_{[1]} - \epsilon)$ even in the worst Bayes Nash equilibrium: reveal no information to bidder j and reveal full information to every other bidder.

The proof of Lemma D.1 follows a similar argument as that of Claim 5.2, thus omitted here. It is important to notice that Lemma D.1 does not depend on the probability δ so long as it is strictly positive. Therefore,

given any realized value profile v, if we can mix it with a value profile u of arbitrarily small probability that satisfied the properties in Lemma D.1, then the corresponding private scheme will force bidder j – the bidder with highest value in value profile u – to bid at least $u_{[2]} = v_{[1]} - \epsilon$. Such a bid will help the auctioneer leverage almost full surplus from v.

For convenience, we will call u the *auxiliary vector* of v. We now show that for any realized value profile $v \in \mathcal{V}$, such an auxiliary vector u can be constructed. We divide the argument into four cases. Recall that we use $\rho^* = \max_{\theta, i} v_i(\theta)$ to denote the maximum value among all bidders' values at all states, and i^* denotes the bidder who has value ρ^*. Similarly, $\rho^{**} = \max_{i \neq i^*; \theta} v_i(\theta)$ is the maximum among all possible valuations, excluding bidder i^*'th, and $i^{**} \neq i^*$ denote a bidder who has the value ρ^{**}.

- **Case 0**: the bidder of the largest value is not unique. Let I denote the set of bidders with the largest value $v_{[1]}$ in profile v. So $|I| > 1$. Construct $w^1, w^2 \in \mathbb{R}^n$ as follows. Let $w^1 = v$; Let $\widetilde{i}$ be the largest index in I and define $w_i^2 = v_i$ for all $i \notin I$, $w_{\widetilde{i}}^2 = v_{\widetilde{i}}$ and $w_i^2 = 0$ for all $i \in I, i \neq \widetilde{i}$. Note that $w^1, w^2 \in \mathcal{V} = \prod_{i=1}^n \mathcal{V}_i$, thus have strictly positive probabilities by our assumption that λ supports on the entire set $\mathcal{V}$. It is easy to verify that a mixture of w^1 with probability $1 - \frac{\epsilon}{v_{[1]}}$ and w^2 with probability $\frac{\epsilon}{v_{[1]}}$ results in an expected value profile u such that $u_i = v_i$ for all $i \notin I$ and $u_{\widetilde{i}} = v_{[1]}$, $u_i = v_{[1]} - \epsilon$ for all $i \in I, i \neq \widetilde{i}$. Such a u satisfies the conditions in Lemma D.1 for any arbitrarily small ϵ. For the following analysis, we assume the vector v has a unique maximum value.

- **Case 1**: $v_{i^*} \leq \rho^{**}$ and v_{i^*} is *not* the largest value in v. Let $\widetilde{i}$ be the bidder with the largest value $v_{[1]}$, thus $v_{\widetilde{i}} = v_{[1]} > v_{i^*}$. Construct $w^1, w^2 \in \mathbb{R}^n$ as follows. Let $w_i^1 = v_i$ for all $i \neq i^*, \widetilde{i}$, and $w_{\widetilde{i}}^1 = v_{[1]}$, $w_{i^*}^1 = \rho^*$; let $w_i^2 = v_i$ for all $i \neq \widetilde{i}, i^*$, and $w_{\widetilde{i}}^2 = 0$, $w_{i^*}^2 = \rho^*$. It is easy to see that a mixture of w^1 with probability $1 - \frac{\epsilon}{v_{[1]}}$ and w^2 with probability $\frac{\epsilon}{v_{[1]}}$ results in an expected value profile u such that $u_i = v_i$ for all $i \neq \widetilde{i}, i^*$ and $u_{\widetilde{i}} = v_{[1]} - \epsilon$, $u_{i^*} = \rho^*$.

- **Case 2**: $v_{i^*} \leq \rho^{**}$ and v_{i^*} is the largest value in v, thus $v_{i^*} = v_{[1]}$. Construct $w^1, w^2 \in \mathbb{R}^n$ as follows. Let $w_i^1 = v_i$ for all $i \neq i^{**}, i^*$, and $w_{i^{**}}^1 = \rho^{**}$, $w_{i^*}^1 = \rho^*$; let $w_i^2 = v_i$ for all $i \neq i^{**}, i^*$, and $w_{i^{**}}^2 = \rho^{**}$, $w_{i^*}^2 = 0$. A mixture of w^1 with probability $\frac{v_{[1]} - \epsilon}{\rho^*}$ and w^2 with probability $\frac{\rho^* - v_{[1]} + \epsilon}{\rho^*}$ results in an expected auxiliary vector u such that $u_i = v_i$

for all $i \neq i^{**}, i^*$ and $u_{i^{**}} = \rho^{**}$, $u_{i^*} = v_{[1]} - \epsilon$.

- **Case 3**: $v_{i^*} > \rho^{**}$, which happens with probability $\sum_{\theta : v_{i^*}(\theta) > \rho^{**}} \lambda_\theta$. This implies that bidder i^* must have the highest value in v. Construct $w^1, w^2 \in \mathbb{R}^n$ as follows. Let $w_i^1 = v_i$ for all $i \neq i^{**}, i^*$, and $w_{i^{**}}^1 = \rho^{**}$, $w_{i^*}^1 = \rho^*$; let $w_i^2 = v_i$ for all $i \neq i^{**}, i^*$, and $w_{i^{**}}^2 = \rho^{**}$, $w_{i^*}^2 = 0$. A mixture of w^1 with probability $\frac{\rho^{**} - \epsilon}{\rho^*}$ and w^2 with probability $\frac{\rho^* - \rho^{**} + \epsilon}{\rho^*}$ results in an expected auxiliary vector u such that $u_i = v_i$ for all $i \neq i^{**}, i^*$ and $u_{i^{**}} = \rho^{**}$, $u_{i^*} = \rho^{**} - \epsilon$. Therefore, the second highest value in u is $\rho^{**} - \epsilon$, which is the revenue at the worst equilibrium (Lemma D.1). Since the full surplus is v_{i^*}, the revenue is at most $v_{i^*} - \rho^{**} + \epsilon$ away from the full surplus. The aggregated surplus loss from this case is thus $\sum_{\theta : v_{i^*}(\theta) > \rho^{**}} \lambda_\theta [v_{i^*}(\theta) - \rho^{**} + \epsilon]$, which is precisely the excluded term in Theorem 5.1, up to an ϵ.

To sum up, the auxiliary vector u can be constructed for any value profile v. Moreover, the construction corresponds to a signaling scheme that mixes two properly chosen value profiles w^1, w^2, as in the above case analysis. We can then use the private signaling scheme in Lemma D.1 – i.e., revealing no information to the bidder of the highest value in the auxiliary vector u and full information to all other bidders – to extract almost full surplus from v. Note that the excluded term in Theorem 5.1 is mainly due to Case 3. Notice that the probability mass of the auxiliary vector u can be arbitrarily small so long as it is positive, therefore the fact that a certain value profile is used for constructing auxiliary vectors has negligible effects when we consider the revenue extracted from that value profile.

The above argument only shows the existence of a private signaling scheme that extracts almost the full surplus. Next, we prove that this scheme can be implemented efficiently. In particular, given any realized value profile $v(\theta)$, we will need to efficiently decide which signal to send with what probability. The main issue here is, when $v(\theta)$ is needed for constructing auxiliary vectors, we have to identify all the value profiles that use $v(\theta)$ to construct their auxiliary vectors, and signal accordingly. Our next lemma shows that this can be done efficiently, thus completing the proof of Theorem 5.1.

Lemma D.2. *The private signaling scheme above can be implemented in* $\mathrm{poly}(n, \sum_i |\mathcal{V}_i|)$ *time.*

Proof. We will show that given any realized value profile, we can explicitly check whether it is used for constructing auxiliary vectors, and if so, of which value

vectors. For convenience, we call the value profiles that are used for constructing auxiliary vectors the *basic vectors*. A key observation is that for any given $v(\theta)$, we can explicitly identify all value profiles (if any) that use $v(\theta)$ to construct their auxiliary vectors, and there are only $\mathrm{poly}(\sum_i |\mathcal{V}_i|)$ of them. This follows a tedious discussion, which essentially reverse the above case analysis. We omit the details here.

With the above observation, we can now easily construct the desired private signaling scheme as follows. Given any realized profile $v(\theta)$, let u be its auxiliary value profile as we constructed and i be the bidder of the highest value in u. If $v(\theta)$ is not a basic vector, we simply reveal $v(\theta)$ to all the bidders except that we tell bidder i the value profile is $v(\theta)$ with probability $1 - \delta$ and u with probability δ, for some preset small enough δ. If $v(\theta)$ is a basic vector. We first find all the ($\mathrm{poly}(\sum_i |\mathcal{V}_i|)$ many) value profiles that use $v(\theta)$ to construct auxiliary vectors. We can compute the probability needed from $v(\theta)$ for each such value profile, and send the signal with corresponding probabilities by referring to the case analysis above. For the left probability mass, we simply treat $v(\theta)$ as a basic vector and signal accordingly. $\square$