Public Signaling in Bayesian Ad Auctions

Francesco Bacchiocchi, **Matteo Castiglioni**, **Alberto Marchesi**, **Giulia Romano** and **Nicola Gatti**
Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133, Milan, Italy
francesco.bacchiocchi@mail.polimi.it
{matteo.castiglioni, alberto.marchesi, giulia.romano, nicola.gatti}@polimi.it

Abstract

We study *signaling* in *Bayesian ad auctions*, in which bidders' valuations depend on a random, unknown state of nature. The auction mechanism has complete knowledge of the actual state of nature, and it can send signals to bidders so as to disclose information about the state and increase revenue. For instance, a state may collectively encode some features of the user that are known to the mechanism only, since the latter has access to data sources unaccessible to the bidders. We study the problem of computing *how the mechanism should send signals to bidders in order to maximize revenue*. While this problem has already been addressed in the easier setting of second-price auctions, to the best of our knowledge, our work is the first to explore ad auctions with more than one slot. In this paper, we focus on *public* signaling and VCG mechanisms, under which bidders truthfully report their valuations. We start with a negative result, showing that, in general, the problem does *not* admit a PTAS unless $\mathsf{P} = \mathsf{NP}$, even when bidders' valuations are known to the mechanism. The rest of the paper is devoted to settings in which such negative result can be circumvented. First, we prove that, with *known valuations*, the problem can indeed be solved in polynomial time when either the number of states d or the number of slots m is fixed. Moreover, in the same setting, we provide an FPTAS for the case in which bidders are *single minded*, but d and m can be arbitrary. Then, we switch to the *random valuations* setting, in which these are randomly drawn according to some probability distribution. In this case, we show that the problem admits an FPTAS, a PTAS, and a QPTAS, when, respectively, d is fixed, m is fixed, and bidders' valuations are bounded away from zero.

1 Introduction

Nowadays, worldwide spending in digital advertising is skyrocketing, and this growth is primarily driven by *ad auctions*. These account for almost all market share, since they are at the core of popular advertising platforms, such as, *e.g.*, those by Google, Amazon, and Facebook. According to a recent report by eMarketer [2021], digital ad spending will reach over $490 billion in 2021 and zoom past half a trillion in 2022.

We study *signaling* in ad auction settings by means of the *Bayesian persuasion* framework [Kamenica and Gentzkow, 2011]. Over the last years, this framework has received considerable attention from the computer science community, due to its applicability to many real-world scenarios, such as, *e.g.*, online advertising [Bro Miltersen and Sheffet, 2012], voting [Alonso and Câmara, 2016; Cheng *et al.*, 2015; Castiglioni *et al.*, 2020a; Castiglioni and Gatti, 2021], traffic routing [Vasserman *et al.*, 2015; Bhaskar *et al.*, 2016; Castiglioni *et al.*, 2021a], recommendation systems [Mansour *et al.*, 2016], security [Rabinovich *et al.*, 2015; Xu *et al.*, 2016], and product marketing [Babichenko and Barman, 2017; Candogan, 2019]. Some recent works also addressed the problem of dealing with uncertain parameters in Bayesian persuasion, boosting its applicability [Castiglioni *et al.*, 2020c; Castiglioni *et al.*, 2021b; Zu *et al.*, 2021; Babichenko *et al.*, 2021; Castiglioni *et al.*, 2022a].

In a standard ad auction, the advertisers (also called bidders) compete for displaying their ads on a limited number of slots, and each bidder has their own private valuation representing how much they value a click on their ad. In this work, we study *Bayesian ad auctions*, which are characterized by the fact that bidders' valuations depend on a random, unknown state of nature. The auction mechanism has complete knowledge of the actual state of nature, and it can send signals to bidders so as to disclose information about the state and increase revenue. In particular, the auction mechanism *commits to a signaling scheme*, which is defined as a randomized mapping from states of nature to signals being sent to the bidders. Our model fits many real-world applications that are *not* captured by classical ad auctions. For instance, a state of nature may collectively encode some features of the user visualizing the ads—such as, *e.g.*, age, gender, or geographical region—that are known to the mechanism only, since the latter has access to data sources unaccessible to the bidders.

We study the problem of computing *a revenue-maximizing signaling scheme for the mechanism*. In particular, we focus on *public* signaling, in which the mechanism can only send a single signal that is observed by all the bidders. Moreover, we restrict our attention to VCG mechanisms, which are widely used in practice and have the appealing property of inducing bidders to truthfully report their valuations. While the signaling problem studied in this paper has already been addressed in the easier setting of second-price auctions, to the best of our knowledge, our work is the first to explore algorithmic signaling in general ad auctions with more than one slot.

1.1 Original Contributions

We start with a negative result, showing that, in general, the revenue-maximizing problem with public signaling does *not* admit a PTAS unless $\mathsf{P} = \mathsf{NP}$, even when bidders' valuations are known to the mechanism. Thus, we address settings in which such a negative result can be circumvented.

First, we show that, in the *known valuations* setting, the problem admits a polynomial-time algorithm when either the number of slots m or the number of states d is fixed. The proposed algorithms work by solving suitably-defined *linear programs* (LPs) of polynomial size, thanks to the crucial property that, when either m or d is fixed, there always exists an optimal signaling scheme using a polynomial number of different signals. Moreover, we also study special instances in which the bidders are *single minded*, but m and d can be arbitrary. In this case, each bidder positively values a click on their ad only when the actual state of nature is a specific (single) state, and all the bidders interested in the same state value a click on their ad for the same amount. By exploiting a particular combinatorial structure of the set of bidders' posterior distributions induced by signaling schemes, we are able to provide an FPTAS in such setting. The algorithm works by applying the ellipsoid method in a non-trivial way, with only access to an approximate polynomial-time separation oracle. The latter is implemented by a rather involved dynamic programming algorithm, which works thanks to the particular structure of the set of bidders' posteriors.

Then, we switch the attention to the *random valuations* setting, where bidders' valuations are unknown to the mechanism, but randomly drawn according to some probability distribution. In this case, we first provide some preliminary results that establish useful connections between the optimal value of the revenue-maximizing problem and that of optimal signaling schemes restricted to suitably-defined finite sets of posterior distributions. These sets are defined so that the expected revenue of the mechanism is *"stable"*, meaning that it does *not* decrease too much when restricting signaling schemes to use posteriors in such sets. In particular, for our results we use sets of q-uniform posteriors, for suitable values of q. As a preliminary step, we also show that it is possible to compute an approximately-optimal signaling scheme having only access to a finite number of samples from the distribution of bidders' valuations. In conclusion, all the preliminary results described so far allow us to prove that, in the random valuations setting, the problem admits an FPTAS, a PTAS, and a QPTAS, when, respectively, d is fixed, m is fixed, and bidders' valuations are bounded away from zero.[1]

1.2 Related Works

To the best of our knowledge, the algorithmic study of signaling in auctions is limited to the *second-price auction*, which can be seen as a special ad auction with a single slot (with the only exception of [Castiglioni *et al.*, 2022b], which studies signaling in posted price auctions).

[Emek *et al.* 2014] study second-price auctions in the known valuations setting. They provide an LP to compute an optimal public signaling scheme. Moreover, they show that it is NP-hard to compute an optimal signaling scheme in the random valuations setting. In our work, we generalize their positive result, in order to provide our polynomial-time algorithm working when the number of slots m is fixed.

[Cheng *et al.* 2015] complement the hardness result of [Emek *et al.*, 2014] by providing a PTAS for the random valuations setting. This result cannot be extended to ad auctions, as we show in our first negative result. However, we provide two generalizations of the result by [Cheng *et al.* 2015]: we provide a PTAS for the random valuations setting with a fixed number of slots m, and a QPTAS when the bidder's valuations are bounded away from zero.

Finally, [Badanidiyuru *et al.* 2018] study algorithms whose running time does *not* depend on the number of states of nature. Moreover, they initiate the study of private signaling schemes, showing that, in second-price auctions, private signaling introduces non-trivial equilibrium selection problems.

In conclusion, there are some works that, while addressing problems different from our problem, are strictly related to it. In particular, [Daskalakis *et al.*, 2016] studies how to jointly optimize the auction mechanism and the signaling scheme, while [Fu *et al.*, 2012; Li and Das, 2019] provide a quantitative analysis on how revealing information increases revenue.

2 Preliminaries

In a standard *ad auction* [Nisan and Ronen, 2001], there is a set $\mathcal{N} := \{1, \dots, n\}$ of *advertisers* (or *bidders*) who compete for displaying their ads on a set $\mathcal{M} := \{1, \dots, m\}$ of *slots*, with $m \leq n$. Each bidder $i \in \mathcal{N}$ is characterized by a *private valuation* $v_i \in [0, 1]$, which represents how much they value a click on their ad. Moreover, each slot $j \in \mathcal{M}$ is associated with a *click through rate* parameter $\lambda_j \in [0, 1]$, which is the probability with which the slot is clicked by a user. W.l.o.g., we assume that the slots are ordered so that $\lambda_1 \geq \ldots \geq \lambda_m$. The auction goes on as follows: first, each bidder $i \in \mathcal{N}$ separately reports a bid $b_i \in [0, 1]$ to the auction mechanism; then, based on the bids, the latter allocates an ad to each slot and defines how much each bidder has to pay the mechanism for a click on their ad. We focus on *truthful* mechanisms, and the VCG mechanism in particular. In truthful mechanisms, allocation and payments are defined so that it is a dominant strategy for each bidder to report their true valuation to the mechanism, namely $b_i = v_i$ for every $i \in \mathcal{N}$. In particular, the allocation implemented by the VCG mechanism orderly assigns the first m bidders in decreasing value of b_i to the first m slots (those with the highest click through rates). At the same time, assuming w.l.o.g. that bidder i is assigned to slot i, the mechanism defines an expected payment $p_i := \sum_{j=i+1}^{m+1} b_j(\lambda_{j-1} - \lambda_j)$ for each bidder $i \in \{1, \dots, m\}$, where, for the ease of notation, we let $\lambda_{m+1} = 0$. The payment is zero for all the other bidders. In practice, each bidder $i \in \{1, \dots, m\}$ has to pay $\frac{p_i}{\lambda_i}$ whenever a user clicks on their

[1] All the proofs are in [Bacchiocchi *et al.*, 2022].

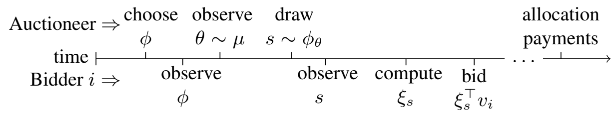

Figure 1: Time-line of a Bayesian ad auction.

ad, so that their utility is $\lambda_i v_i - p_i$ in expectation over the clicks. The expected utility of all the other bidders is zero. We study *Bayesian* ad auctions, which are characterized by a set $\Theta := \{\theta_1, \dots, \theta_d\}$ of d states of nature. Each bidder $i \in \mathcal{N}$ has a valuation vector $v_i \in [0,1]^d$, with $v_i(\theta)$ being bidder i's valuation in state $\theta \in \Theta$, and all such vectors are arranged in a matrix of bidders' valuations $V \in [0,1]^{n \times d}$, whose entries are defined as $V(i,\theta) := v_i(\theta)$ for all $i \in \mathcal{N}$ and $\theta \in \Theta$. We model signaling by means of the *Bayesian persuasion* framework [Kamenica and Gentzkow, 2011]. We consider the case in which the auction mechanism knows the state of nature and acts as a *sender* by issuing signals to the bidders (the *receivers*), so as to partially disclose information about the state and increase revenue. As customary, we assume that the state is drawn from a common prior distribution $\mu \in \Delta_\Theta$, with μ_θ being the probability of $\theta \in \Theta$.[2] The mechanism publicly *commits to a signaling scheme* ϕ, which is a randomized mapping from states of nature to signals for the bidders. We focus on the *public signaling* case, where all the bidders receive the same signal. Formally, a signaling scheme is a function $\phi : \Theta \to \Delta_{\mathcal{S}}$, where $\mathcal{S}$ is a set of available signals. For the ease of notation, we let $\phi_\theta(s)$ be the probability of sending signal $s \in \mathcal{S}$ when the state is $\theta \in \Theta$.

A Bayesian ad auction goes on as follows (Figure 1): (i) the auction mechanism commits to a signaling scheme ϕ, and the bidders observe it; (ii) the mechanism gets to know the state of nature $\theta \sim \mu$ and draws signal $s \sim \phi(\theta)$; and (iv) the bidders observe the signal s and rationally update their prior belief over states according to Bayes rule. After observing $s \in \mathcal{S}$, all the bidders infer a *posterior* distribution $\xi_s \in \Delta_\Theta$ over state such that the posterior probability of state $\theta \in \Theta$ is

$$\xi_s(\theta) := \frac{\mu_\theta \phi_\theta(s)}{\sum_{\theta' \in \Theta} \mu_{\theta'} \phi_{\theta'}(s)}. \tag{1}$$

Finally, each bidder $i \in \mathcal{N}$ truthfully reports to the mechanism their expected valuation given the posterior ξ_s, namely $\xi_s^\top v_i = \sum_{\theta \in \Theta} v_i(\theta)\, \xi_s(\theta)$, and the mechanism allocates slots and defines payments as in a standard ad auction.[3]

Representing Signaling Schemes. It is oftentimes useful to represent signaling schemes as convex combinations of the posteriors they can induce [Dughmi, 2014; Cheng *et al.*, 2015]. Formally, a signaling scheme $\phi : \Theta \to \Delta_{\mathcal{S}}$ *induces* a probability distribution γ over posteriors in Δ_Θ, with $\gamma(\xi)$ denoting the probability of posterior $\xi \in \Delta_\Theta$, defined as

$$\gamma(\xi) := \sum_{s \in \mathcal{S} : \xi_s = \xi} \sum_{\theta \in \Theta} \mu_\theta \phi_\theta(s).$$

[2]Given a finite set X, we denote with Δ_X the $(|X| - 1)$-dimensional simplex defined over the elements of X.

[3]Given that all the bidders share the same posterior, the induced auction is equivalent to a classical VCG ad auction. Thus, truthfully reporting expected valuations is a dominant strategy for the bidders.

Indeed, we can directly reason about distributions γ over Δ_Θ rather than about signaling schemes, provided that they are *consistent* with the prior. By letting $\text{supp}(\gamma) := \{\xi \in \Delta_\Theta \mid \gamma(\xi) > 0\}$ be the support of γ, this requires that

$$\sum_{\xi \in \text{supp}(\gamma)} \gamma(\xi)\, \xi(\theta) = \mu_\theta \quad \forall \theta \in \Theta. \tag{2}$$

In the rest of the paper, we will use the term signaling scheme to refer to a consistent distribution γ over Δ_Θ.

Computational Problems. We focus on the problem of computing an *optimal* signaling scheme, *i.e.*, one maximizing the revenue of the mechanism. We study two settings:

- the *known valuations* (KV) setting in which the matrix of bidders' valuations V is known to the mechanism; and
- the *random valuations* (RV) setting in which the matrix of bidders' valuations V is unknown, but randomly drawn according to a probability distribution $\mathcal{V}$.

As it is customary in the literature (see, *e.g.*, [Badanidiyuru *et al.*, 2018]), in the RV setting we assume that algorithms have access to a black-box oracle returning i.i.d. samples drawn from $\mathcal{V}$ (rather than actually knowing such distribution). We denote by $\textsc{Rev}(V,\xi)$ the expected revenue of the mechanism when the bidders' valuations are given by V and the posterior induced by the mechanism is $\xi \in \Delta_\Theta$. Formally, given that bidders truthfully report their expected valuations and assuming w.l.o.g. that bidder i is assigned by the mechanism to slot i, we can write $\textsc{Rev}(V,\xi) := \sum_{j=1}^m j\, \xi^\top v_{j+1}(\lambda_j - \lambda_{j+1})$. Then, given a signaling scheme γ, the expected revenue of the mechanism is $\sum_{\xi \in \text{supp}(\gamma)} \gamma(\xi)\, \textsc{Rev}(V,\xi)$. When the valuations are unknown, we let $\textsc{Rev}(\mathcal{V},\xi) := \mathbb{E}_{V \sim \mathcal{V}} \textsc{Rev}(V,\xi)$ and define the expected revenue analogously. Notice that, given a distribution of valuations $\mathcal{V}$ (or, in the KV setting, a matrix of bidders' valuations V) and a finite set $\Xi \subseteq \Delta_\Theta$ of posteriors, it is possible to formulate the problem of computing an optimal signaling scheme as an LP, as follows:[4]

$$\max_{\gamma \in \Delta_\Xi} \sum_{\xi \in \Xi} \gamma(\xi)\, \textsc{Rev}(\mathcal{V},\xi) \quad \text{s.t.} \tag{3a}$$

$$\sum_{\xi \in \Xi} \gamma(\xi)\, \xi(\theta) = \mu_\theta \qquad \forall \theta \in \Theta. \tag{3b}$$

In the following, we let OPT_Ξ be the optimal value of LP 3, while we denote with OPT the optimal expected revenue of the mechanism over all the possible signaling schemes γ.[5]

3 A General Inapproximability Result

We start our analysis with the following negative result:

Theorem 1. *The problem of computing an optimal signaling scheme does* not *admit a* PTAS *unless* P = NP*, even when it is restricted to the KV setting.*

[4]LP 3 is written for the RV setting, its analogous for the KV setting can be obtained by substituting $\textsc{Rev}(\mathcal{V},\xi)$ with $\textsc{Rev}(V,\xi)$.

[5]The dependence of OPT_Ξ and OPT from either V or $\mathcal{V}$ is omitted, as it will be clear from context.

Theorem 1 is proved by a reduction from the VERTEX COVER problem in cubic graphs [Alimonti and Kann, 2000]. In the rest of this work, we study several settings in which the negative result in Theorem 1 can be circumvented, by either fixing some parameters of the problem (see Sections 4 and 6.1) or considering instances with a specific structure (see Sections 5 and 6.2).

4 KV Setting: Parametrized Complexity

In this section, we study the parametrized complexity of the problem of computing an optimal signaling scheme, showing that it admits a polynomial-time algorithm when either the number of slots m or the number of states of nature d is fixed.

In the following, we let $\Pi_l \subseteq 2^{\mathcal{N}}$ be the set of all the the possible permutations of $l \leq n$ bidders taken from $\mathcal{N}$, with $\pi = (i_1, ..., i_l) \in \Pi_l$ denoting a tuple made by bidders $i_1, \ldots, i_l \in \mathcal{N}$, in that order. We also let $\Xi_\pi \subseteq \Delta_\Theta$ be the (possibly empty) polytope of posteriors in which the expected valuations of bidders in $\pi \in \Pi_l$ are ordered (from the highest to the lowest) according to π; formally, it holds $\Xi_\pi := \left\{\xi \in \Delta_\Theta \mid \xi^\top v_{i_1} \geq \xi^\top v_{i_2} \geq \ldots \geq \xi^\top v_{i_l}\right\}$. Notice that, given a permutation $\pi \in \Pi_l$ of $l \geq m+1$ bidders, the expected revenue of the mechanism in any posterior $\xi \in \Xi_\pi$ is $\text{REV}(V, \xi) := \sum_{j=1}^{m} j\, \xi^\top v_{j+1}(\lambda_j - \lambda_{j+1})$, since the bidders truthfully report their expected valuations to the mechanism, and, thus, the latter allocates slots to bidders in π according to their order in the permutation. Thus, for any fixed $\pi \in \Pi_l$ with $l \geq m + 1$, the term $\text{REV}(V, \xi)$ is linear in ξ over Ξ_π.

4.1 Fixing the Number of Slots m

In this case, the problem can be solved in polynomial time by formulating it as an LP, thanks to the following lemma:

Lemma 1. *There always exists an optimal signaling scheme γ such that $|\Xi_\pi \cap \text{supp}(\gamma)| \leq 1$ for every $\pi \in \Pi_{m+1}$.*

Intuitively, the lemma follows from the fact that, given any signaling scheme γ and two posteriors $\xi, \xi' \in \text{supp}(\gamma)$ such that $\xi, \xi' \in \Xi_\pi$ for some $\pi \in \Pi_{m+1}$, it is always possible to define a new signaling scheme that replaces ξ and ξ' with a suitably-defined convex combination of them, without decreasing the expected revenue (since it is linear over Ξ_π).

By Lemma 1, we can re-write the revenue maximization problem as $\max \sum_{\pi \in \Pi_{m+1}} \gamma(\xi_\pi)\text{REV}(V, \xi_\pi)$ subject to constraints ensuring that each ξ_π belongs to Ξ_π (for $\pi \in \Pi_{m+1}$) and that γ is a consistent probability distribution over such posteriors (see Equation (2)). This problem can be formulated as an LP by introducing a variable for each $\pi \in \Pi_{m+1}$ and $\theta \in \Theta$, encoding the products $\gamma(\xi_\pi)\xi_\pi(\theta)$ that define the expected revenue. Overall, the resulting LP (see LP 8 in [Bacchiocchi *et al.*, 2022]) has a number of variables and constraints that is $O(n^m)$, which, after fixing m, is polynomial in the size of the input.[6] Thus, we conclude that:

Theorem 2. *In the KV setting, if the number of slots m is fixed, then an optimal signaling scheme can be computed in polynomial time.*

[6] We remark that LP 8 in [Bacchiocchi *et al.*, 2022] is a generalization of the LP presented by Emek *et al.* [2014] for the easier case of second price auctions.

4.2 Fixing the Number of States d

Our polynomial-time algorithm exploits the fact that an optimal signaling scheme can be computed by restricting the attention to distributions supported on a finite set of posteriors whose cardinality is polynomial in all the parameters, except from d. In particular, it is sufficient to focus on the set $\Xi^* := \bigcup_{\pi \in \Pi_n} V(\Xi_\pi)$, where $V(\cdot)$ denotes the set of vertices of the polytope given as input. Formally:

Lemma 2. *It holds that $OPT_{\Xi^*} = OPT$.*

The lemma follows from the fact that, given any signaling scheme γ and posterior $\xi \in \text{supp}(\gamma)$ such that $\xi \in \Xi_\pi$ for some $\pi \in \Pi_n$, by Carathéodory's theorem it is always possible (since Ξ_π is a polytope) to decompose ξ into a convex combination of the vertices of Ξ_π, obtaining a new signaling scheme that provides the mechanism with an expected revenue at least as large as that of γ (since $\text{REV}(V, \xi)$ is linear over Ξ_π). By observing that $|\Xi^*| = O((n^2 + d)^{d-1})$, it is easy to show that an optimal signaling scheme can be computed by means of LP 3 instantiated for the set Ξ^*, which has a number of variables and constraints that is polynomial once d is fixed. This proves the following:

Theorem 3. *In the KV setting, if the number of states d is fixed, then an optimal signaling scheme can be computed in polynomial time.*

5 KV Setting: Single-Minded Bidders

In this section, we focus on particular Bayesian ad auctions where the bidders are *single minded*. Intuitively, in our setting, by single mindedness we mean that each bidder is interested in displaying their ad only when the realized state of nature is a specific (single) state, and that all the bidders interested in the same state value a click on their ad for the same amount. We introduce the following formal definition:

Definition 1 (Single-minded bidders)**.** *In a Bayesian ad auction, we say that bidders are* single minded *if there exist $\mathcal{N}_\theta \subseteq \mathcal{N}$ and $\delta_\theta \in [0, 1]$ for all $\theta \in \Theta$ such that:*

(i) $\mathcal{N} = \bigcup_{\theta \in \Theta} \mathcal{N}_\theta$ and $\mathcal{N}_\theta \cap \mathcal{N}_{\theta'} = \emptyset$ for all $\theta \neq \theta' \in \Theta$;

(ii) for every $\theta \in \Theta$ and $i \in \mathcal{N}_\theta$, it holds $v_i(\theta) = \delta_\theta$ and $v_i(\theta') = 0$ for all $\theta' \in \Theta : \theta' \neq \theta$.

Notice that, given a posterior $\xi \in \Delta_\Theta$ induced by the mechanism, all the bidders i belonging to the same set $\mathcal{N}_\theta$ have the same expected valuation, namely $\xi^\top v_i = \delta_\theta \xi(\theta)$ for all $\theta \in \Theta$ and $i \in \mathcal{N}_\theta$. As a result, given that bidders truthfully report their expected valuations, the mechanism will always receive at most d different bids, one per set $\mathcal{N}_\theta$.

The last observation implies that, given $\xi \in \Delta_\Theta$, in order to unequivocally define an allocation of bidders to slots (and, thus, also define the expected payments) it is sufficient to know the relative ordering of the (at most) d different expected valuations associated to sets $\mathcal{N}_\theta$. This allows us to tackle the problem with an approach analogous to the one of Section 4, with the only difference that, in this case, we will reason about permutations of the groups of bidders $\mathcal{N}_\theta$, rather than about permutations of all the individual bidders.

In the following, we let $\Pi \subseteq 2^\Theta$ be the set of all the permutations of the sates of nature $\Theta = \{\theta_1, \ldots, \theta_d\}$, while we let

$\pi = (\theta_{k_1}, \dots, \theta_{k_d}) \in \Pi$ be an ordered tuple made by states $\theta_{k_1}, \dots, \theta_{k_d} \in \Theta$, where $k_1, \dots, k_d \in \{1, \dots, d\}$. Moreover, $\Xi_\pi := \left\{ \xi \in \Delta_\Theta \mid \delta_{\theta_{k_1}} \xi(\theta_{k_1}) \geq \dots \geq \delta_{\theta_{k_d}} \xi(\theta_{k_d}) \right\}$ is the polytope of posteriors in which the expected valuations associated to sets $\mathcal{N}_\theta$ are ordered according to π.

The first preliminary result that we need in order to derive our approximation algorithm is a characterization of the vertices of the sets Ξ_π for $\pi \in \Pi$, as follows.

Lemma 3. *Given $\pi \in \Pi$ and $\xi \in \Xi_\pi$, it holds that $\xi \in V(\Xi_\pi)$ if and only if there exists $\ell \in \{1, \dots, d\}$ such that:*

(i) $\delta_{\theta_{k_1}} \xi(\theta_{k_1}) = \dots = \delta_{\theta_{k_\ell}} \xi(\theta_{k_\ell}) > 0$*; and*

(ii) $\delta_{\theta_{k_{\ell+1}}} \xi(\theta_{k_{\ell+1}}) = \dots = \delta_{\theta_{k_d}} \xi(\theta_{k_d}) = 0.$

Intuitively, Lemma 3 states that the vertices of a set Ξ_π are all the posteriors $\xi \in \Delta_\Theta$ such that, for some $\ell \in \{1, \dots, d\}$, only the first ℓ states according to the ordering defined by π are assigned a positive probability, while all the remaining states have zero probability. Moreover, the positive probabilities of the posterior ξ are defined so that all the bidders belonging to the first ℓ sets $\mathcal{N}_\theta$, according to the ordering defined by π, are the same. Notice that, in the special case in which all the values δ_θ are equal to one, the vertices of all the sets Ξ_π are all the uniform probability distributions over subsets of ℓ states of nature, for any $\ell \in \{1, \dots, d\}$.

By letting $\Xi^* = \bigcup_{\pi \in \Pi} V(\Xi_\pi)$, since the term $\textsc{Rev}(V, \xi)$ is linear in ξ over Ξ_π for every permutation $\pi \in \Pi$, we can conclude that $OPT_{\Xi^*} = OPT$ (the proof is analogous to that of Lemma 2). Thus, Lemma 3 allows us to find an optimal signaling scheme by solving LP 3 for the set Ξ^* and the matrix of bidders' valuations V. However, notice that, since the size of Ξ^* is exponential in d, the resulting LP has exponentially-many variables. Nevertheless, since the LP has polynomially-many constraints, we can still solve it in polynomial time by applying the ellipsoid algorithm to its dual, provided that a polynomial-time separation oracle is available.

In order to design a polynomial-time separation oracle, we apply the procedure described above to a relaxed version of LP 3, whose optimal value is sufficiently "close" to that of the original LP. Given $\beta \in \mathbb{R}_+$, the relaxed LP reads as follows:

$$\max_{\gamma \in \Delta_{\Xi^*}, z \leq 0} \sum_{\xi \in \Xi} \gamma(\xi) \, \textsc{Rev}(\mathcal{V}, \xi) + \beta z \quad \text{s.t.} \tag{4a}$$

$$\sum_{\xi \in \Xi^*} \gamma(\xi) \, \xi(\theta) - z \geq \mu_\theta \qquad \forall \theta \in \Theta. \tag{4b}$$

The dual problem of LP 4 reads as follows:

$$\min_{y \leq 0, t} \sum_{\theta \in \Theta} y_\theta \mu_\theta + t \quad \text{s.t.} \tag{5a}$$

$$\sum_{\theta \in \Theta} y_\theta \, \xi(\theta) + t \geq \textsc{Rev}(\mathcal{V}, \xi) \qquad \forall \xi \in \Xi^* \tag{5b}$$

$$\sum_{\theta \in \Theta} y_\theta \geq -\beta, \tag{5c}$$

where y_θ for $\theta \in \Theta$ are dual variables associated to Constraints (3b), while t is a dual variable for $\sum_{\xi \in \Xi^*} \gamma(\xi) = 1$. Notice that, by relaxing the LP, in the dual LP 5 we get the additional Constraint (5c) and that $y_\theta \leq 0$ for all $\theta \in \Theta$. This is crucial to design a polynomial-time separation oracle.

The separation problem associated to Problem 5 reads as:

Definition 2 (Separation problem)**.** *Given values for the dual variables $y_\theta \in [-\beta, 0]$ for all $\theta \in \Theta$, compute:*

$$\max_{\xi \in \Xi^*} \textsc{Rev}(V, \xi) - \sum_{\theta \in \Theta} y_\theta \, \xi(\theta). \tag{6}$$

The following Lemma 4 shows that Problem 6 can be solved optimally up to any given additive loss $\lambda > 0$, by means of a *dynamic programming* algorithm that runs in time polynomial in the size of the input, in $\frac{1}{\lambda}$, and in β. Formally:

Lemma 4. *Given $\lambda > 0$, there exists an algorithm that finds an additive λ-approximation to Problem 6, in time polynomial in the size of the input, in $\frac{1}{\lambda}$, and in β.*

The crucial observation that allows us to solve Problem 6 by means of dynamic programming is that, in any posterior $\xi \in \Xi^*$, bidders' expected valuations are either a positive, bidder-independent value or zero (see Lemma 3). This allows us to build a discretized range of possible bidders' valuation values, so that, for each discretized value, we can compute an optimal posterior $\xi \in \Xi^*$ inducing that value by adding states of nature incrementally in a dynamic programming fashion.

Since the algorithm in Lemma 4 only returns an approximate solution to Problem 6, we need to carefully apply the ellipsoid algorithm to solve LP 5, so that it correctly works even with an approximated oracle. Some non-trivial duality arguments allow us to prove that, indeed, this can be achieved by only incurring in a small additive loss on the quality of the returned solution, and without degrading the running time of the algorithm. Overall, this allows us to conclude that:

Theorem 4. *In the KV setting, if the bidders are single minded, then the problem of computing an optimal signaling scheme admits an (additive) FPTAS.*

6 RV Setting

In this setting, as stated in Section 2, we assume that the auction mechanism has access to the distribution of bidders' valuations $\mathcal{V}$ only through a black-box sampling oracle. In the following, given $s \in \mathbb{N}_{>0}$ i.i.d. samples of matrices of bidders' valuations, namely $V_1, \dots, V_s \in [0, 1]^{n \times d}$, we let $\mathcal{V}^s$ be their empirical distribution, which is such that $\Pr_{V \sim \mathcal{V}^s} \left\{ V = \hat{V} \right\} := \frac{\sum_{t=1}^s \mathbb{1}\{V_t = \hat{V}\}}{s}$ for all $\hat{V} \in [0, 1]^{n \times d}$.

In this section, we first study the parametrized complexity of the problem of computing an optimal signaling scheme in general auctions (Section 6.1), and, then, we address special auction settings in which the bidders' valuations are *bounded away from zero*, namely $v_i(\theta) > \delta$ for all $i \in N$ and $\theta \in \Theta$, for some threshold $\delta > 0$. In the latter case, we show that the problem admits a QPTAS and the result is tight (Section 6.2).

Before stating our main results (Theorems 5, 6, 7, and 8), we introduce some preliminary useful lemmas. The first one (Lemma 5) works under the true distribution of bidders' valuations $\mathcal{V}$, and it establishes a connection between the optimal expected revenue (OPT) and the optimal value of LP 3 for suitably-defined finite sets $\Xi \subseteq \Delta_\Theta$ of posteriors (OPT_Ξ).

In particular, we look at sets $\Xi \subseteq \Delta_\Theta$ for which the function $\text{REV}(\mathcal{V}, \cdot)$ is "*stable*" according to the following definition:[7]

Definition 3 ((α, ε)-stability). *Given $\alpha, \varepsilon \geq 0$ and a finite set $\Xi \subseteq \Delta_\Theta$, we say that $\text{REV}(\mathcal{V}, \cdot)$ is (α, ε)-stable for Ξ if, for every $\xi \in \Delta_\Theta$, there exists a distribution $\gamma_\xi \in \Delta_\Xi$ such that:*

$$\sum_{\xi' \in \Xi} \gamma_\xi(\xi')\text{REV}(\mathcal{V}, \xi') \geq (1-\alpha)\text{REV}(\mathcal{V}, \xi) - \varepsilon. \quad (7)$$

For any finite set $\Xi \subseteq \Delta_\Theta$ such that $\text{REV}(\mathcal{V}, \cdot)$ is (α, ε)-*stable* for Ξ, starting from an optimal signaling scheme γ one can recover an optimal solution to LP 3, only incurring in "small" multiplicative and additive losses in the expected revenue, respectively of $1-\alpha$ and ε. This can be accomplished by decomposing each posterior $\xi \in \text{supp}(\gamma)$ into $\gamma_\xi \in \Delta_\Xi$ and, then, putting such distributions together. These observations allow us to prove the following lemma:

Lemma 5. *Given $\alpha, \varepsilon \geq 0$ and $\Xi \subseteq \Delta_\Theta$ such that $\text{REV}(\mathcal{V}, \cdot)$ is (α, ε)-stable for Ξ, it holds $OPT_\Xi \geq (1-\alpha)OPT - \varepsilon$.*

The second lemma (Lemma 6) deals with the approximation error introduced by using an empirical distribution of bidders' valuations $\mathcal{V}^s$, rather than the actual distribution $\mathcal{V}$. Given a finite set $\Xi \subseteq \Delta_\Theta$ of posteriors, let $\gamma_{\mathcal{V}^s} \in \Delta_\Xi$ be an optimal solution to LP 3 for distribution $\mathcal{V}^s$ and set Ξ. Moreover, let $OPT_{\Xi,s} := \mathbb{E}\left[\sum_{\xi \in \Xi} \gamma_{\mathcal{V}^s}(\xi)\text{REV}(\mathcal{V}, \xi)\right]$ be the average expected revenue of signaling schemes $\gamma_{\mathcal{V}^s}$ under the true distribution of valuations $\mathcal{V}$, where the expectation is with respect to the sampling procedure that determines $\mathcal{V}^s$. Then, a concentration argument proves the following:

Lemma 6. *Given $\rho, \tau > 0$, let $\Xi \subseteq \Delta_\Theta$ be finite and $s := \left\lceil \frac{2(\lambda_1 m)^2}{\tau^2} \log \frac{2}{\rho} \right\rceil$, $OPT_{\Xi,s} \geq (1-\rho|\Xi|)\, OPT_\Xi - \tau$.*

Finally, the last lemma (Lemma 7) exploits Lemma 5 to provide two useful bounds on the value of OPT_{Ξ_q}, where $\Xi_q \subseteq \Delta_\Theta$ (for a given $q \in \mathbb{N}_{>0}$) is the finite set of all the q-uniform posteriors, according to the following definition:

Definition 4 (q-uniform posterior). *Given $q \in \mathbb{N}_{>0}$, a posterior $\xi \in \Delta_\Theta$ is q-uniform if each $\xi(\theta)$ is a multiple of $\frac{1}{q}$.*

Notice that the set Ξ_q has size $|\Xi^q| \leq \min\{d^q, q^d\}$. The two points in the following lemma are readily proved by applying Lemma 5, after noticing that the sets Ξ_q in the statement are such that the function $\text{REV}(\mathcal{V}, \cdot)$ is (α, ε)-stable for them, with suitable values of $\alpha \geq 0$ and $\varepsilon \geq 0$. Formally:

Lemma 7. *Given $\eta > 0$ and $q := \left\lceil \frac{1}{2\eta^2} \log \frac{m+1}{\eta} \right\rceil$, it holds:*

(i) $OPT_{\Xi_q} \geq OPT - 2\eta m$;

(ii) if, for some $\delta > 0$, it is the case that $v_i(\theta) > \delta$ for all $i \in \mathcal{N}$ and $\theta \in \Theta$, then $OPT_{\Xi_q} \geq (1 - \frac{\eta}{\delta})^2\, OPT$.

6.1 Parametrized Complexity

First, we study the computational complexity of the problem of computing an optimal signaling scheme when the number of states d is fixed. We provide an (additive) FPTAS that works by performing the following two steps: (i) it collects a suitable number $s \in \mathbb{N}_{>0}$ of matrices of bidders' valuations, by invoking the sampling oracle; and (ii) it solves LP 3 for the resulting empirical distribution $\mathcal{V}^s$ and a suitably-defined set of q-uniform posteriors. In particular, given a desired (additive) error $\lambda > 0$, the algorithm works on the set Ξ_q for $q = \lceil \frac{md}{\lambda} \rceil$ and its approximation guarantees rely on the following Lemma 8, proved again by means of Lemma 5.

Lemma 8. *For $\lambda > 0$ and $q = \lceil \frac{md}{\lambda} \rceil$, $OPT_{\Xi_q} \geq OPT - \lambda$.*

Thanks to Lemmas 6 and 8 (the former applied for suitable values $\rho, \tau > 0$), we can prove that the procedure described in steps (i) and (ii) above gives a signaling scheme achieving an expected revenue at most a function of λ lower than OPT, provided that the number of samples s is defined as in Lemma 8. Moreover, let us notice that, since $|\Xi_q| = O(q^d) = O((\frac{1}{\lambda} md)^d)$, if d is fixed, then the overall procedure runs in time polynomial in the input size and in $\frac{1}{\lambda}$. Thus, we can conclude that:

Theorem 5. *In the RV setting, if the number of states d is fixed, then the problem of computing an optimal signaling scheme admits and (additive) FPTAS.*

Next, we switch the attention to the case in which the number of slots m is fixed. We provide an (additive) PTAS that works as the FPTAS in Theorem 5, but whose approximation guarantees follow from Lemma 6 and point (i) in Lemma 7 (rather than Lemma 8). Thus, the only difference with respect to the previous case is that the algorithm works on the set Ξ_q of q-uniform posteriors for q defined as in Lemma 7. As a result, since $|\Xi_q| = O(d^q)$ and q depends on a parameter $\eta > 0$ that is related to the quality of the obtained approximation, the algorithm is only a PTAS rather than an FPTAS. Formally, we can prove the following:

Theorem 6. *In the RV setting, if the number of slots m is fixed, then the problem of computing an optimal signaling scheme admits and (additive) PTAS.*

6.2 Valuations Bounded Away From Zero

We conclude the section by studying the case in which the bidders' valuations are bounded away from zero. This case is dealt with an algorithm identical to the one in Theorem 6, but carrying on the approximation analysis by using Lemma 6 and point (ii) in Lemma 7 (rater than point (i)). Thus, since the value of q in Lemma 7 is related to the quality of the approximation thorough a parameter $\eta > 0$ and also depends logarithmically on the number of slots m, we obtain:

Theorem 7. *In the RV setting, if $v_i(\theta) \geq \delta$ for all $i \in \mathcal{N}$ and $\theta \in \Theta$ for some $\delta > 0$, then the problem of computing an optimal signaling scheme admits a (multiplicative) QPTAS.*

The following theorem shows that the result is tight.

Theorem 8. *Assuming the ETH, there exists a constant $\omega > 0$ such that finding a signaling scheme that provides an expected revenue at least of $(1-\omega)OPT$ requires $I^{\tilde{\Omega}(\log I)}$ time, where I is the size of the problem instance. This holds even when $v_i(\theta) > \frac{1}{3}$ for all $i \in \mathcal{N}$ and $\theta \in \Theta$.*[8]

[7]Notions of stability analogous to that in Definition 3 have already been used in the literature; see, *e.g.*, [Cheng *et al.*, 2015].

[8]The $\tilde{\Omega}$ notation hides poly-logarithmic factors.

References

[Alimonti and Kann, 2000] Paola Alimonti and Viggo Kann. Some APX-completeness results for cubic graphs. *Theoretical Computer Science*, 237:123–134, 04 2000.

[Alonso and Câmara, 2016] Ricardo Alonso and Odilon Câmara. Persuading voters. *American Economic Review*, 106(11):3590–3605, 2016.

[Babichenko and Barman, 2017] Yakov Babichenko and Siddharth Barman. Algorithmic aspects of private Bayesian persuasion. In *ITCS*, 2017.

[Babichenko *et al.*, 2021] Yakov Babichenko, Inbal Talgam-Cohen, Haifeng Xu, and Konstantin Zabarnyi. Regret-minimizing bayesian persuasion. In *EC*, page 128, 2021.

[Bacchiocchi *et al.*, 2022] Francesco Bacchiocchi, Matteo Castiglioni, Alberto Marchesi, Giulia Romano, and Nicola Gatti. Public signaling in bayesian ad auctions. *CoRR*, abs/2201.09728, 2022.

[Badanidiyuru *et al.*, 2018] Ashwinkumar Badanidiyuru, Kshipra Bhawalkar, and Haifeng Xu. Targeting and signaling in ad auctions. In *SODA*, pages 2545–2563, 2018.

[Bhaskar *et al.*, 2016] Umang Bhaskar, Yu Cheng, Young Kun Ko, and Chaitanya Swamy. Hardness results for signaling in bayesian zero-sum and network routing games. In *EC*, pages 479–496, 2016.

[Bro Miltersen and Sheffet, 2012] Peter Bro Miltersen and Or Sheffet. Send mixed signals: earn more, work less. In *EC*, pages 234–247, 2012.

[Candogan, 2019] Ozan Candogan. Persuasion in networks: Public signals and k-cores. In *EC*, pages 133–134, 2019.

[Castiglioni and Gatti, 2021] Matteo Castiglioni and Nicola Gatti. Persuading voters in district-based elections. In *AAAI*, pages 5244–5251, 2021.

[Castiglioni *et al.*, 2020a] Matteo Castiglioni, Andrea Celli, and Nicola Gatti. Persuading voters: It's easy to whisper, it's hard to speak loud. In *AAAI*, pages 1870–1877, 2020.

[Castiglioni *et al.*, 2020b] Matteo Castiglioni, Andrea Celli, and Nicola Gatti. Public bayesian persuasion: Being almost optimal and almost persuasive. *CoRR*, abs/2002.05156, 2020.

[Castiglioni *et al.*, 2020c] Matteo Castiglioni, Andrea Celli, Alberto Marchesi, and Nicola Gatti. Online Bayesian persuasion. In *NeurIPS*, pages 16188–16198, 2020.

[Castiglioni *et al.*, 2021a] Matteo Castiglioni, Andrea Celli, Alberto Marchesi, and Nicola Gatti. Signaling in Bayesian network congestion games: the subtle power of symmetry. In *AAAI*, pages 5252–5259, 2021.

[Castiglioni *et al.*, 2021b] Matteo Castiglioni, Alberto Marchesi, Andrea Celli, and Nicola Gatti. Multi-receiver online Bayesian persuasion. In *ICML*, pages 1314–1323, 2021.

[Castiglioni *et al.*, 2022a] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Bayesian persuasion meets mechanism design: Going beyond intractability with type reporting. In *AAMAS*, 2022.

[Castiglioni *et al.*, 2022b] Matteo Castiglioni, Giulia Romano, Alberto Marchesi, and Nicola Gatti. Signaling in posted price auctions. In *AAAI*, 2022.

[Cheng *et al.*, 2015] Yu Cheng, Ho Yee Cheung, Shaddin Dughmi, Ehsan Emamjomeh-Zadeh, Li Han, and Shang-Hua Teng. Mixture selection, mechanism design, and signaling. In *FOCS*, pages 1426–1445, 2015.

[Daskalakis *et al.*, 2016] Constantinos Daskalakis, Christos H. Papadimitriou, and Christos Tzamos. Does information revelation improve revenue? In *EC*, pages 233–250, 2016.

[Dughmi, 2014] Shaddin Dughmi. On the hardness of signaling. In *FOCS*, pages 354–363. IEEE, 2014.

[eMarketer, 2021] eMarketer. Worldwide digital ad spending year-end update. https://www.emarketer.com/content/worldwide-digital-ad-spending-year-end-update, 2021. Accessed: 2022-01-09.

[Emek *et al.*, 2014] Yuval Emek, Michal Feldman, Iftah Gamzu, Renato PaesLeme, and Moshe Tennenholtz. Signaling schemes for revenue maximization. *TEAC*, 2(2):1–19, 2014.

[Fu *et al.*, 2012] Hu Fu, Patrick R. Jordan, Mohammad Mahdian, Uri Nadav, Inbal Talgam-Cohen, and Sergei Vassilvitskii. Ad auctions with data. In *SAGT*, pages 168–179, 2012.

[Kamenica and Gentzkow, 2011] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[Li and Das, 2019] Zhuoshu Li and Sanmay Das. Revenue enhancement via asymmetric signaling in interdependent-value auctions. In *AAAI*, pages 2093–2100, 2019.

[Mansour *et al.*, 2016] Yishay Mansour, Aleksandrs Slivkins, Vasilis Syrgkanis, and Zhiwei Steven Wu. Bayesian exploration: Incentivizing exploration in Bayesian games. In *EC*, pages 661–661, 2016.

[Nisan and Ronen, 2001] Noam Nisan and Amir Ronen. Algorithmic mechanism design. *Games and Economic behavior*, 35(1-2):166–196, 2001.

[Rabinovich *et al.*, 2015] Zinovi Rabinovich, Albert Xin Jiang, Manish Jain, and Haifeng Xu. Information disclosure as a means to security. In *AAMAS*, pages 645–653, 2015.

[Vasserman *et al.*, 2015] Shoshana Vasserman, Michal Feldman, and Avinatan Hassidim. Implementing the wisdom of waze. In *IJCAI*, pages 660–666, 2015.

[Xu *et al.*, 2016] Haifeng Xu, Rupert Freeman, Vincent Conitzer, Shaddin Dughmi, and Milind Tambe. Signaling in Bayesian Stackelberg games. In *AAMAS*, pages 150–158, 2016.

[Zu *et al.*, 2021] You Zu, Krishnamurthy Iyer, and Haifeng Xu. Learning to persuade on the fly: Robustness against ignorance. In *EC*, pages 927–928, 2021.