Public Bayesian Persuasion: Being Almost Optimal and Almost Persuasive

Matteo Castiglioni[1] · **Andrea Celli**[2] · **Nicola Gatti**[1]

Abstract

We study *algorithmic Bayesian persuasion* problems in which the principal (a.k.a. the *sender*) has to persuade multiple agents (a.k.a. *receivers*) by using public communication channels. Specifically, our model follows the multi-receiver model with no inter-agent externalities introduced by Arieli and Babichenko (J Econ Theory 182:185–217, 2019). It is known that the problem of computing a sender-optimal public persuasive signaling scheme is not approximable even in simple settings. Therefore, prior works usually focus on determining restricted classes of the problem for which efficient approximation is possible. Typically, positive results in this space amounts to finding *bi-criteria* approximation algorithms yielding an almost optimal and almost persuasive solution in polynomial time. In this paper, we take a different perspective and study the persuasion problem in the general setting where the space of the states of nature, the action space of the receivers, and the utility function of the sender can be arbitrary. We fully characterize the computational complexity of computing a bi-criteria approximation of an optimal public signaling scheme in such settings. In particular, we show that, assuming the Exponential Time Hypothesis, solving this problem requires at least a quasi-polynomial number of steps even in instances with simple utility functions and binary action spaces such as an election with the k-voting rule. In doing so, we prove that a relaxed version of the MAXIMUM FEASIBLE SUBSYSTEM OF LINEAR INEQUALITIES problem requires at least quasi-polynomial time to be solved. Finally, we close the gap by providing a quasi-polynomial time bi-criteria approximation algorithm for arbitrary public persuasion problems that, under mild assumptions, yields a QPTAS.

✉ Matteo Castiglioni
   matteo.castiglioni@polimi.it

   Andrea Celli
   andrea.celli2@unibocconi.it

   Nicola Gatti
   nicola.gatti@polimi.it

[1] Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy

[2] Bocconi University, Via Roberto Sarfatti 25, 20136 Milan, Italy

1 Introduction

In many real-life strategic interactions, agents rely on information revealed by an exogenous entity to take decisions. The latter acts as an informed *principal* whose goal is to shape the agents' beliefs so as to achieve a desired outcome. In this context, deciding what information to reveal amounts to an *information structure design* problem. When information is incomplete, the information structure determines "which agent gets to know what" about the current state of the environment (i.e., the parameters determining payoff functions). There has been a recent surge of interest in the study of how an informed principal may steer agents' collective behavior towards a favorable outcome. The study of these problems has been largely driven by their application in various domains such as auctions and online advertisement [1–3], voting [4, 5], traffic routing [6, 7], recommendation systems [8], security [9–11], and marketing [12, 13].

Persuasion is the task faced by an informed principal, which we call the *sender*, which tries to influence the behavior of the self-interested agent(s) (i.e., the *receivers*) taking part in a strategic interactions. The sender faces the algorithmic problem of determining the optimal information structure to achieve her objectives. A solution to this problem is described through the notion of *signaling scheme*, which is a mapping from the sender's observations to the space of probability distributions over the set of available signals. A foundational model describing the persuasion problem is the *Bayesian persuasion framework* (BP) introduced by Kamenica and Gentzkow in [14]. That model describes a setting with a sender and a single receiver. There is a set of parameters influencing the payoff functions of the sender and of the receiver. These parameters are collectively called the *state of nature*, and model exogenous stochasticity in the environment. The sender and the receiver share a common prior over the possible states of nature. However, only the sender gets to observe the realized state of nature, which is drawn according to the shared prior probability distribution. This originates a fundamental asymmetry in the information available to the two agents. The sender can exploit this additional knowledge to steer the receiver's actions towards a favorable outcome. Specifically, the action selected by the receiver is the best action available under her current posterior distribution, which is updated in a classical Bayesian fashion after observing the sender's signal. Therefore, the prior distribution together with the sender's signaling scheme determine the receiver equilibrium behavior. We observe that the BP framework assumes the sender's commitment power, which is a natural assumption in many settings (see, e.g., the arguments by [14, 15]). One argument to that effect is that reputation and credibility may be a key factor for the long-term utility of the sender [16].

In many practical scenarios, the sender may need to persuade *multiple receivers*, revealing information to each one of them. In the multi-receiver setting, it is useful

to make a distinction between *private* and *public* signaling schemes.[1] In the former setting, the sender may reveal different information to each receiver through private communication channels. In the latter, which is the focus of this paper, the sender has to reveal the same information to all receivers. Public persuasion is well suited for settings where private communication channels are either too costly or impractical. This is the case in scenarios with a large population of receivers, such as elections, and scenarios where receivers may share private information, which are frequent in practice.

In our paper, we adopt and generalize the multi-agent persuasion model introduced by Arieli and Babichenko in [18], which rules out the possibility of *inter-agent externalities*. Specifically, each receiver's utility depends only on her action and on the realized state of nature, but not on the actions of other receivers. This assumption allows one to focus on the key problem of coordinating the receivers' behavior, without the additional complexity arising from externalities which have been shown to make the problem largely intractable [6, 19]. Previous works on Arieli and Babichenko's model either address the private persuasion setting [18, 20, 21] or make some structural assumptions which render them special cases of our model [22]. To the best of our knowledge, this is the first work which generalizes the model by Arieli and Babichenko to settings with arbitrary spaces of states of nature, arbitrary receivers' action spaces, and arbitrary sender's utility functions. The generality of our setting raises a number of technical difficulties with respect to previous works on the same model. Our solution to these challenges is a first step towards actionable persuasion models that can be applied to real-world multi-receiver problems without structural restrictions.

1.1 Context: Persuasion with Multiple Receivers

Dughmi and Xu [23] analyze for the first time Bayesian persuasion from a computational perspective, focusing on the single receiver case. In [18], Arieli and Babichenko introduce the model of persuasion with multiple receivers and without inter-agent externalities, with a focus on private Bayesian persuasion. In particular, they study the setting with a binary action space for the receivers and a binary space of states of nature. They provide a characterization of the optimal signaling scheme in the case of supermodular, anonymous submodular, and supermajority sender's utility functions. In [20], Babichenko and Barman extend the work by Arieli and Babichenko providing a tight $(1 - 1/e)$-approximate signaling scheme for monotone submodular sender's utilities and showing that an optimal private scheme for anonymous utility functions can be found efficiently. In [21], Dughmi and Xu generalize the previous model to settings with an arbitrary number of states of nature.

When considering the problem of designing public persuasive signaling schemes, some previous works study scenarios with inter-agent externalities by making some structural assumptions on the nature of the strategic interaction. For instance, Bhaskar et al. [6] and Rubinstein [19] study public signaling problems in which two receivers

---

[1] For the sake of completeness, we also mention that some recent work studies signaling schemes in the middle between private and public, called *semi-public* [17].

play a zero-sum game. In particular, Bhaskar et al. rule out an additive PTAS assuming hardness of the planted clique problem. The setting studied by Bhaskar et al. [6] and Rubinstein [19] is fundamentally different from ours. In their setting the game can be compactly represented through its normal-form representation, and the complexity of the problem lies in handling externalities among players. On the other hand, in our setting, a compact normal-form representation is not possible since it is exponential in the (arbitrary) number of receivers.

Moreover, Rubinstein proves that the problem of computing an ϵ-optimal signaling scheme requires at least quasi-polynomial time assuming the Exponential Time Hypothesis (ETH). This result is tight due to the quasi-polynomial approximation scheme proposed by Cheng et al. [5].

A number of previous works focus on the public signaling problem in the no inter-agent externalities framework of Arieli and Babichenko. In particular, Dughmi and Xu [21] rule out the existence of a PTAS even when receivers have binary action spaces and objectives are linear, unless P = NP. For this reason, most of the following works focus on the computation of bi- or tri-criteria approximations in which the persuasion constraints can be violated by a small amount. In [5], Cheng et al. describe a polynomial-time tri-criteria approximation algorithm for k-voting scenarios. The work of [5] on k-voting is related to the voting problem that we study in this paper. However, while we relax the problem allowing approximately optimal and approximately persuasive signaling schemes, [5] considers also a third type of relaxation. In particular, they consider a relaxed sender's utility function in which less than k votes are sufficient to win the election. This third relaxation is necessary to provide a PTAS to the problem, while we show that without relaxing the utility function the problem requires at least quasi-polynomial time. In [22], Xu studies public persuasion with binary action spaces and an arbitrary number of states of nature, showing that no bi-criteria FPTAS is possible, unless P = NP. Furthermore, the author proposes a bi-criteria PTAS for monotone submodular sender's utility functions and shows that, when the number of states of nature is fixed and a non-degeneracy assumption holds, an optimal signaling scheme can be computed in polynomial time.

1.2 Our Results and Techniques

We provide a *tight* characterization of the complexity of computing bi-criteria approximations of optimal public signaling schemes in arbitrary persuasion problems with n receivers and no inter-agent externalities.

Impossibility result Previous works studying the same model (i.e., one with no inter-agent externalities and public signaling schemes) exploit specific structures of the sender's utility functions to provide optimal or approximate polynomial-time algorithms. We show that the complexity of the approximation problem shifts from *poly-time* to *quasi-poly-time* when the utility function of the sender can be arbitrary. Indeed, we show that the positive results by Xu [22], which assumes *noise-stability* of the sender's utility function, cannot be extended to the case of arbitrary sender's utility functions. Specifically, we argue that no polynomial-time bi-criteria approximation algorithm is possible in general settings. This is shown by reasoning over simple

k-voting instances, which are per se an interesting application scenario of Bayesian persuasion (see, e.g., the work by Castiglioni et al. [24]), and are sufficient to extend the result to the general case of arbitrary sender's utility functions. In addressing this result, we follow a different approach from that used in [22]. Specifically, we cannot hope for a 'standard' NP-hardness result because there exist quasi-polynomial time bi-criteria approximation algorithms (Theorem 2). Therefore, by assuming ETH, we show that it is unlikely that there exists a bi-criteria polynomial-time approximation algorithm even in instances with simple utility functions and a fixed space of actions. Let n be the size of the instance in input. Our main impossibility result reads as follows.

Theorem 1 *Assuming ETH, there exists a constant $\epsilon^* > 0$ such that, for any $0 < \epsilon \le \epsilon^*$, finding a signaling scheme that is ϵ-persuasive and α-approximate requires time $n^{\tilde{\Omega}(\log n)}$ for any multiplicative or additive factor $\alpha \in (0, 1)$, even with binary action spaces.*

The proof of this result requires an intermediate step that is of independent interest and of general applicability. Specifically, we study a slight variation of the MAXIMUM FEASIBLE SUBSYSTEM OF LINEAR INEQUALITIES problem (ϵ-MFS) [5], where, given a linear system $A\mathbf{x} \ge 0$, $A \in [-1, 1]^{n_{\mathrm{row}} \times n_{\mathrm{col}}}$, we look for the vector $\mathbf{x} \in \Delta_{n_{\mathrm{col}}}$ almost (i.e., except for an additive factor ϵ) satisfying the highest number of inequalities (Definition 4). This is a constrained variant of the MAX FLS problem previously studied by Amaldi and Kann in [25], and it is commonly used in scheduling [26], signaling, and mechanism design [5]. In Sect. 5, we prove that solving ϵ-MFS requires at least a quasi-polynomial number of steps assuming ETH. The proof is based on a reduction from *two-provers games* [27, 28]. Then, equipped with the result on ϵ-MFS, we focus on a simple public persuasion problem where the receivers are voters, and they have a binary action space since they must choose one between two candidates. In Sect. 6, we prove a hardness result (Theorem 8) for this setting which directly implies Theorem 1. We show that the ϵ-MFS problem is deeply connected to the problem of computing 'good' posteriors, as finding of an optimal $\mathbf{x}$ in ϵ-MFS maps to the problem of finding an ϵ-persuasive posterior, which is equivalent to determining an ϵ-persuasive signaling scheme.

Positive result In order to design an approximation algorithm (in the multiplicative sense), we resort to the assumption of α-approximable utility functions for the sender, as previously defined by Xu in [22]. An α-approximable sender's utility function is such that it is possible to obtain in polynomial time a tie-breaking for the receivers guaranteeing to the sender an α-approximation of the optimal objective value. The α-approximability condition is a natural minimal requirement since, otherwise, even the problem of evaluating the sender's objective function for a given posterior over the states of nature would not be tractable. When the sender's utility function is α-approximable, there is no hope for a better approximation than an α-approximate signaling scheme. The following theorem, presented in Sect. 7, shows that it is possible to compute, in quasi-polynomial time, a bi-criteria approximation with a factor arbitrarily close to α.

Theorem 2 *Let the sender's utility function f be α-approximable. Then, for any $\delta > 0$ and $\epsilon > 0$, there exists a poly $\left(n^{\log(n/\delta)/\epsilon^2} \right)$ algorithm that outputs an $\alpha(1-\delta)$-approximate ϵ-persuasive public signaling scheme.*

Therefore, our approximation algorithm guarantees the best possible factor on the objective value, and an arbitrary small loss in persuasiveness. For 1-approximable functions, Theorem 2 yields a bi-criteria QPTAS. In the setting of Xu [22] (i.e., binary action spaces and state-independent sender's utility function), our result directly yields a QPTAS for any monotone sender's utility function. In order to prove the result, we show that any posterior can be represented as a convex combination of k-uniform posteriors with only a small loss in the objective value. By restricting our attention to the set of k-uniform posteriors, which has a quasi-polynomial size, the problem can be solved via a linear program of quasi-polynomial size.

2 Preliminaries

This section describes the instantiation of the Bayesian persuasion framework which is the focus of this work (Sect. 2.1), public signaling problems (Sect. 2.2) and the notion of the bi-criteria approximation which we adopt (Sect. 2.3). For a comprehensive overview of the Bayesian persuasion framework we refer the reader to [15], [29], and [30].[2]

2.1 Basic Model

Our model is a generalization of the framework introduced by Arieli and Babichenko in [18], that is, multi-agent persuasion with no inter-agent externalities. We adopt the perspective of a *sender* facing a finite set of *receivers* $\mathcal{R} := [\bar{n}]$. Each receiver $r \in \mathcal{R}$ has a finite set of ϱ^r actions $\mathcal{A}^r := \{a_i\}_{i=1}^{\varrho^r}$. Each receiver's payoff depends only on the action she takes and on a (random) *state of nature* θ, drawn from a finite set $\Theta := \{\theta_i\}_{i=1}^{d}$ of cardinality d. In particular, receiver r's utility is given by the function $u^r : \mathcal{A}^r \times \Theta \to [0, 1]$. The utility of each receiver does not depend on other receivers' actions because of the no inter-agent externalities assumption. We denote by $u_\theta^r(a^r) \in [0, 1]$ the utility observed by receiver r when the state of nature is θ and she plays a^r. Let $\mathcal{A} := \times_{r \in \mathcal{R}} \mathcal{A}^r$ be the set of joint receivers' actions. An action profile (i.e., a tuple specifying an action for each receiver) is denoted by $\mathbf{a} = (a^r)_{r=1}^{\bar{n}} \in \mathcal{A}$. The sender's utility, when the state of nature is θ, is given by the function $f_\theta : \mathcal{A} \to [0, 1]$. We write $f_\theta(\mathbf{a})$ to denote the sender's payoff when the receivers behave according to action profile $\mathbf{a}$ and the state of nature is θ. As it is customary in Bayesian persuasion, we assume f_θ can be represented succinctly, that is without explicitly describing the function through its (exponentially many) input–output pairs. As an example, the reader can refer to Eq. 3, where it is possible

---

[2] Throughout the paper, the set $\{1, \ldots, x\}$ is denoted by $[x]$, $\text{int}(X)$ is the *interior* of set X, and Δ_X is the set of all probability distributions over X. The indicator function for the set $\mathcal{E}$ is denoted by $I[\mathcal{E}]$. Bold case letters denote column vectors. Moreover, we generally denote the size of a problem input by n.

to compute the sender's payoff by reasoning on the structure of the action profile at hand.

The state of nature θ is drawn from a common prior distribution $\mu \in \text{int}(\Delta_\Theta)$, which is explicitly known to the sender and the receivers. Moreover, the sender can publicly commit to a policy ϕ (i.e., a *signaling scheme*, see Sect. 2.2) which maps states of nature to *signals* for the receivers. A generic signal for receiver r is denoted by s^r, while the set of available signals to each receiver r is denoted by $\mathcal{S}^r$. The interaction between the sender and the receivers goes as follows:

1. The sender commits to a publicly known signaling scheme ϕ;
2. The sender observes the realized state of nature $\theta \sim \mu$;
3. The sender draws a signal s^r for each receiver according to the signaling scheme ϕ_θ, and communicates to each receiver r the signal s^r;
4. Each receiver r observes s^r and updates her prior beliefs over Θ following Bayes rule. Then, each receiver r selects an action $a^r \in \mathcal{A}^r$ maximizing her expected reward.

Let $\mathbf{a} = (a^1, \ldots, a^{\bar{n}}) \in \mathcal{A}$ be the tuple of the receivers' choices, then each receiver r gets payoff $u_\theta^r(a^r)$, and the sender observes payoff $f_\theta(\mathbf{a})$. This work focuses on the specific setting in which ϕ is a *public signaling schemes*. We give more details on the structure of public signaling schemes in the following section.

2.2 Public Signaling Schemes

A *signal profile* is a tuple $\mathbf{s} = (s^r)_{r=1}^{\bar{n}} \in \mathcal{S}$ specifying a signal for each receiver, where $\mathcal{S} := \times_{r \in \mathcal{R}} \mathcal{S}^r$. A *public signaling scheme* is a function $\phi : \Theta \to \Delta_\mathcal{S}$ mapping states of nature to probability distributions over signal profiles, with the constraint that each receiver has to receive the same signal, that is, for any θ and $\mathbf{s} \sim \phi_\theta$, it holds $s^r = s^{r'}$ for each pair of receivers r, r'. With an overload of notation we write $s \in \mathcal{S}$ to denote the public signal received by all receivers. The probability with which the sender selects s after observing θ is denoted by $\phi_\theta(s)$. Thus, it holds $\sum_{s \in \mathcal{S}} \phi_\theta(s) = 1$ for each $\theta \in \Theta$.

After observing $s \in \mathcal{S}$, receiver r performs a Bayesian update and infers a posterior belief $\mathbf{p} \in \Delta_\Theta$ over the states of nature. Specifically, the realized state of nature is θ with probability

$$p_\theta := \frac{\mu_\theta \, \phi_\theta(s)}{\sum_{\theta \in \Theta} \mu_\theta \, \phi_\theta(s)}.$$

Since the prior is common knowledge and all receivers observe the same s, they all perform the same Bayesian update and have the same posterior belief regarding the realized state of nature. After computing $\mathbf{p}$, since the problem is without inter-agent externalities, each receiver solves a disjoint single-agent decision problem to find the action maximizing her expected utility.

A signaling scheme is *direct* when signals can be mapped to actions of the receivers, and interpreted as action recommendations. Each receiver is sent the same signal $\mathbf{s} \in \mathcal{A}$ specifying a (possibly different) action for each other receiver, that is, the set

of possible signals is $\mathcal{S} = \mathcal{A}$. Moreover, a signaling scheme is *persuasive* if following the recommendations is an equilibrium of the underlying *Bayesian game* [31, 32]. A direct signaling scheme is persuasive if the sender's action recommendations belongs to $\arg\max_{a \in \mathcal{A}^r} \sum_{\theta \in \Theta} p_\theta \, u_\theta^r(a)$ for every receiver r. A simple revelation-principle style argument shows that there always exists an optimal public signaling scheme which is both direct and persuasive [17, 22]. A signal in a direct signaling scheme can be equivalently expressed as an action profile $\mathbf{a} \in \mathcal{A}$. It is easy to see that there is an exponential number of such signals. We write $\phi_\theta(\mathbf{a})$ to denote the probability with which the sender selects $\mathbf{s} = \mathbf{a}$ when the realized state of nature is θ. The problem of determining an optimal public signaling scheme which is direct and persuasive can be formulated with the following (exponentially sized) linear program (LP):

$$\max \sum_{\theta \in \Theta, \mathbf{a} \in \mathcal{A}} \mu_\theta \, \phi_\theta(\mathbf{a}) \, f_\theta(\mathbf{a}) \tag{1a}$$

$$\text{s.t.} \sum_{\theta \in \Theta} \mu_\theta \, \phi_\theta(\mathbf{a}) \left(u_\theta^r(a^r) - u_\theta^r(a') \right) \geq 0 \qquad \forall r \in \mathcal{R}, \forall \mathbf{a} \in \mathcal{A}, a' \in \mathcal{A}^r$$

$$\sum_{\mathbf{a} \in \mathcal{A}} \phi_\theta(\mathbf{a}) = 1 \qquad\qquad\qquad \forall \theta \in \Theta$$

$$\phi_\theta(\mathbf{a}) \geq 0 \qquad\qquad\qquad \forall \theta \in \Theta, \forall \mathbf{a} \in \mathcal{A} \tag{1b}$$

where $\mathbf{a} = (a^r)_{r=1}^{\bar{n}} \in \mathcal{A}$. The sender's goal is computing the signaling scheme maximizing her expected utility (Objective Function (1a)). Constraints (1b) force the public signaling scheme to be persuasive.[3,4]

2.3 Bi-criteria Approximation

Let $\epsilon \in [0, 1]$. We say that a public signaling scheme is ϵ-*persuasive* if the following holds for any $r \in \mathcal{R}$, $\mathbf{a} \in \mathcal{A}$, and $a' \in \mathcal{A}^r$:

$$\sum_{\theta \in \Theta} \mu_\theta \, \phi_\theta(\mathbf{a}) \left(u_\theta^r(a^r) - u_\theta^r(a') \right) \geq -\epsilon. \tag{2}$$

Throughout the paper, we focus on the computation of approximately optimal signaling schemes. Let OPT be the optimal value of LP (1), i.e., the best expected revenue

---

[3] The interaction between the sender and receivers could be modeled by introducing a single *meta*-receiver with action set $\mathcal{A}$, that is, an action set equal to the Cartesian product of the action sets of the original receivers. However, from a computational perspective, this would not help in assessing the tractability of the problem. In particular, the new *meta*-receiver has an *exponential* action space (being the Cartesian product of the original sets). Our negative result (Theorem 1) shows that, in this setting, concavification approaches analogous to what proposed for the single receiver case by Kamenica and Gentzkow [14] would not be tractable.

[4] Note that the persuasiveness Constraints (1b) are different from the ones defining private signaling schemes. Indeed, since each receiver observes all the action recommendations in public signaling schemes, in LP (1) there is a constraint for each *tuple* of actions/signal $\mathbf{a}$, receiver r, and deviation a'. In private signaling schemes, each receiver observes only her action recommendation and hence there is a persuasiveness constraint for each *action*/signal a, receiver r, and deviation a'.

that the sender can reach under public persuasion constraints. For each state of nature θ, f_θ is a non-negative function, and we have that OPT ≥ 0. When a signaling scheme yields an expected sender utility of at least α OPT, with $\alpha \in (0, 1]$, we say that the signaling scheme is α-*approximate* (that is, approximate in multiplicative sense). When a signaling scheme yields an expected sender utility of at least OPT $- \alpha$, with $\alpha \in [0, 1)$, we say that the scheme is α-*optimal* (that is, approximate in additive sense).

Finally, we consider approximations which relax both the optimality and the persuasiveness constraints. When a signaling scheme is both ϵ-persuasive and α-approximate (or α-optimal), we say it is a *bi-criteria approximation*. We say that one such signaling scheme is (α, ϵ)-*persuasive*.

3 An Application: Persuasion in Voting Problems

In order to clarify the framework we just described, we present a simple example of a possible application of public Bayesian persuasion with no inter-agent externalities. This example is going to be useful in the remainder of the paper.

In an election with the k-voting rule, candidates are elected if they receive at least $k \in [\bar{n}]$ votes (see Brandt et al. [33] for further details). In this setting, a sender (e.g., a politician or a lobbyist) may send signals to the voters on the basis of her private information which is hidden to them. After observing the sender's signal, each voter (i.e., one of the receivers) chooses one among the set of candidates.

In the following, we will employ instances of k-voting in which receivers have to choose one between two candidates. Then, they have a binary action space with actions a_0 and a_1 corresponding to choosing the first or the second candidate, respectively. Each receiver r has utility $u_\theta^r(a) \in [0, 1]$ for each $a \in \{a_0, a_1\}$, where $\theta \in \Theta$. The sender's preferred candidate is the one corresponding to action a_0. Therefore, her objective is maximizing the probability that a_0 receives more than k votes. Formally, the sender's utility function is such that $f_\theta = f$ for each θ, and

$$f(\mathbf{a}) := \begin{cases} 1 & \text{if } |\{r \in \mathcal{R} : a^r = a_0\}| \geq k \\ 0 & \text{otherwise} \end{cases} \qquad \text{for each } \mathbf{a} \in \mathcal{A}. \qquad (3)$$

Moreover, let $W : \Delta_\Theta \to \mathbb{N}_0^+$ be a function returning, for a given posterior distribution $\mathbf{p} \in \Delta_\Theta$, the number of receivers such that $\sum_\theta p_\theta (u_\theta^r(a_0) - u_\theta^r(a_1)) \geq 0$, i.e., the number of voters that will vote for a_0 with a persuasive signaling scheme. Analogously, $W_\epsilon(\mathbf{p})$ is the number of receivers for which $\sum_\theta p_\theta (u_\theta^r(a_0) - u_\theta^r(a_1)) \geq -\epsilon$, i.e., the number of voters that will vote for a_0 with an ϵ-persuasive signaling scheme. In the above voting setting, we refer to the problem of finding an ϵ-persuasive signaling scheme which is also α-approximate (or α-optimal) as (α, ϵ)-k-voting. To further clarify this election scenario, we provide the following simple example, adapted from Castiglioni et al. [24].

Example 1 There are three voters $\mathcal{R} = \{1, 2, 3\}$ who must select one between two candidates $\{a_0, a_1\}$. The sender (e.g., a politician or a lobbyist) observes the realized state of nature, drawn from the uniform probability distribution $(1/3, 1/3, 1/3)$ over

Table 1 Voters' payoffs from voting different candidates

		State A		State B		State C	
		a_0	a_1	a_0	a_1	a_0	a_1
Voters	1	$+1$	$-1/4$	-1	$-1/4$	-1	$-1/4$
	2	-1	$-1/4$	$+1$	$-1/4$	-1	$-1/4$
	3	-1	$-1/4$	-1	$-1/4$	$+1$	$-1/4$

Table 2 Optimal signaling scheme

		Signals		
		not A	not B	not C
States	A	0	1/2	1/2
	B	1/2	0	1/2
	C	1/2	1/2	0

$\Theta = \{A, B, C\}$, and exploits this information to support the election of a_0. The state of nature describes the position of a_0 on a matter of particular interest to the voters. Moreover, all the voters have a slightly negative opinion of candidate a_1, independently of the state of nature, while the opinion on candidate a_0 can be better or worse than the opinion on a_1 depending the state of nature. Table 1 describes the utility of the three voters.

We consider a k-voting rule with $k = 2$. Without any form of signaling, all the voters would vote for a_1 because it provides an expected utility of $-1/4$, against $-1/3$, and the sender would get a utility of 0. If the sender discloses all the information regarding the state of nature (i.e., with a fully informative signal), the sender would still get a utility of 0, since two out of three receivers would pick a_1 in each of the possible states. However, the sender can design a public signaling scheme guaranteeing herself a utility of 1 for each state of nature. Table 2 describes one such scheme with arbitrary signals. Suppose the observed state is A, and that the signal sent by the sender is not B. Then, the posterior distribution over the states of nature is $(1/2, 0, 1/2)$. Therefore, receiver 1 and receiver 3 would vote for a_0 since their expected utility would be 0 against $-1/4$. Similarly, for any other signal, two receivers vote for a_0. Then, the sender's expected payoff is 1. We can recover an equivalent direct signaling scheme by sending a tuple with a candidates' suggestion for each voter. For example, not A would become (a_1, a_0, a_0), and each voter would observe the recommendations given to the others.

4 Technical Toolkit

In this section, we summarize some key results previously studied in the literature that we will extensively use in the remainder of the paper. In particular, we describe some of the results on two-prover games by Babichenko et al. [34] and Deligkas et al. [28] (Sect. 4.1), and we describe a useful theorem on error-correcting codes due to Gilbert [35] (Sect. 4.2).

4.1 Two-Prover Games

A *two-prover* game $\mathcal{G}$ is a co-operative game played by two players (Merlin_1 and Merlin_2, respectively), and an adjudicator (verifier) called Arthur. At the beginning of the game, Arthur draws a pair of questions $(x, y) \in \mathcal{X} \times \mathcal{Y}$ according to a probability distribution $\mathcal{D}$ over the joint set of questions (i.e., $\mathcal{D} \in \Delta_{\mathcal{X} \times \mathcal{Y}}$). Merlin_1 (resp., Merlin_2) observes x (resp., y) and chooses an answer ξ_1 (resp., ξ_2) from her finite set of answers Ξ_1 (resp., Ξ_2). Then, Arthur declares the Merlins to have *won* with a probability equal to the value of a *verification function* $\mathcal{V}(x, y, \xi_1, \xi_2)$. A *strategy* for Merlin_1 is a function $\eta_1 : \mathcal{X} \to \Xi_1$ mapping each possible question to an answer. Analogously, $\eta_2 : \mathcal{Y} \to \Xi_2$ is a strategy of Merlin_2. Before the beginning of the game, Merlin_1 and Merlin_2 can agree on their pair of (possibly mixed) strategies (η_1, η_2), but no communication is allowed during the games. The payoff of a game $\mathcal{G}$ to the Merlins under (η_1, η_2) is defined as: $u(\mathcal{G}, \eta_1, \eta_2) := \mathbb{E}_{(x,y) \sim \mathcal{D}}[\mathcal{V}(x, y, \eta_1(x), \eta_2(y))]$. The *value* of a two-prover game $\mathcal{G}$, denoted by $\omega(\mathcal{G})$, is the maximum expected payoff to the Merlins when they play optimally: $\omega(\mathcal{G}) := \max_{\eta_1} \max_{\eta_2} u(\mathcal{G}, \eta_1, \eta_2)$. The size of the game is $|\mathcal{G}| = |\mathcal{X} \times \mathcal{Y} \times \Xi_1 \times \Xi_2|$.

A two-prover game is called a *free game* if $\mathcal{D}$ is a uniform probability distribution over $\mathcal{X} \times \mathcal{Y}$. This implies that there is no correlation between the questions sent to Merlin_1 and Merlin_2. It is possible to build a family of free games mapping to 3SAT formulas arising from Dinur's PCP theorem. We say that the size n of a formula φ is the number of variables plus the number of clauses in the formula. Moreover, $\mathrm{SAT}(\varphi) \in [0, 1]$ is the maximum fraction of clauses that can be satisfied in φ. With this notation, the Dinur's PCP Theorem reads as follows:

Theorem 3 *[Dinur's PCP Theorem [36]] Given any 3SAT instance φ of size n, and a constant $\rho \in (0, 1/8)$, we can produce in polynomial time a 3SAT instance φ' such that:*

1. *The size of φ' is $n \cdot polylog(n)$;*
2. *Each clause of φ' contains exactly 3 variables, and every variable is contained in at most $d = O(1)$ clauses;*
3. *If $SAT(\varphi) = 1$, then $SAT(\varphi') = 1$;*
4. *If $SAT(\varphi) < 1$, then $SAT(\varphi') < 1 - \rho$.*

A 3SAT formula can be seen as a bipartite graph in which the left vertices are the variables, the right vertices are the clauses, and there is an edge between a variable and a clause whenever that variable appears in that clause. Then, such a bipartite graph has constant degree since each vertex has constant degree. This holds because each clause has at most 3 variables and each variable is contained in at most d clauses. A useful result on bipartite graphs is the following.

Lemma 1 *(Lemma 1 of Deligkas et al. [28]) Let (V, E) be a bipartite graph with $|V| = n$, where U and W are the two disjoints and independent sets such that $V = U \cup W$ (i.e., U and W are the two sides of the graph), and where each vertex has a degree at most v.*

Suppose that U and W both have a constant fraction of the vertices, i.e., $|U| = c \cdot n$ and $|W| = (1 - c) \cdot n$ for some $c \in [0, 1)$.

Then, we can efficiently find a partition $\{S_i\}_{i=1}^{\sqrt{n}}$ of U, and a partition $\{T_j\}_{j=1}^{\sqrt{n}}$ of W, such that each set has a size of at most $2\sqrt{n}$, and for all i and j we have $|(S_i \times T_j) \cap E| \le 2\nu^2$.

Lemma 1 can be used to build the following free game.

Definition 1 (Definition 2 of Deligkas et al. [28]) Given a 3SAT formula φ of size n, we define a free game $\mathcal{F}_\varphi$ as follows:

1. Arthur applies Theorem 3 to obtain formula φ' of size $n \cdot \text{polylog}(n)$;
2. let $m = \sqrt{n \cdot \text{polylog}(n)}$. Arthur applies Lemma 1 to partition the variables of φ' in sets $\{S_i\}_{i=1}^m$, and the clauses in sets $\{T_j\}_{j=1}^m$;
3. Arthur draws an index i uniformly at random from $[m]$, and draws independently an index j uniformly at random from $[m]$. Then, he sends S_i to Merlin$_1$ and T_j to Merlin$_2$;
4. Merlin$_1$ responds by choosing a truth assignment for each variable in S_i, and Merlin$_2$ responds by choosing a truth assignment to every variable that is involved with a clause in T_j;
5. Arthur awards the Merlins payoff 1 if and only if the following conditions are both satisfied:

   - Merlin$_2$'s assignment satisfies all clauses in T_j;
   - the two Merlins' assignments are *compatible*, i.e., for each variable v appearing in S_i and each clause in T_j that contains v, Merlin$_1$'s assignment to v agrees with Merlin$_2$'s assignment to v;

Arthur awards payoff 0 otherwise.

When computing the Merlins' rewards, the second condition is always satisfied when S_i and T_j share no variables. Moreover, when Merlin$_1$'s and Merlin$_2$'s assignments are not compatible, we say that they are *in conflict*. The following lemma shows that, if φ is unsatisfiable, then the value of the corresponding free game $\mathcal{F}_\varphi$ is bounded away from 1.

Lemma 2 *(Lemma 2 by Deligkas et al. [28]) Given a 3SAT formula φ, the following holds:*

- *If φ is satisfiable then $\omega(\mathcal{F}_\varphi) = 1$;*
- *If φ is unsatisfiable then $\omega(\mathcal{F}_\varphi) \le 1 - \rho/2\nu$.*

We define FREEGAME$_\delta$ as a specific problem within the class of *promise problems* (see, e.g., Even et al. [37], Goldreich [38]).

Definition 2 [FREEGAME$_\delta$] A FREEGAME$_\delta$ problem is defined as:

- `INPUT`: a free game $\mathcal{F}_\varphi$ and a constant $\delta \in (0, 1)$.
- `OUTPUT`: YES-instances: $\omega(\mathcal{F}_\varphi) = 1$; No-instances: $\omega(\mathcal{F}_\varphi) \le 1 - \delta$.

Finally, we will need to assume the *Exponential Time Hypothesis* (ETH), which conjectures that any deterministic algorithm solving 3SAT requires $2^{\Omega(n)}$ time.

Theorem 4 *[Theorem 2 by Deligkas et al. [28]] Assuming ETH, there exists a constant $\delta = \rho/2\nu$ such that* FREEGAME$_\delta$ *requires time $n^{\tilde{\Omega}(\log n)}$.*[5]

---

[5] As usual, we use the notation $\tilde{\Omega}$ to suppress polylogarithmic factors.

4.2 Error-Correcting Codes

A *message* of length $k \in \mathbb{N}_+$ is encoded as a *block* of length $n \in \mathbb{N}_+$, with $n \geq k$. A *code* is a mapping $e : \{0, 1\}^k \to \{0, 1\}^n$. Moreover, let $\text{dist}(e(x), e(y))$ be the *relative Hamming distance* between $e(x)$ and $e(y)$, which is defined as the Hamming distance weighted by $1/n$. The *rate* of a code is defined as $R := k/n$. Finally, the *relative distance* $\text{dist}(e)$ of a code e is the maximum value d^{REL} such that $\text{dist}(e(x), e(y)) \geq d^{\text{REL}}$ for each $x, y \in \{0, 1\}^k$.

In the following, we will need an infinite sequence of codes $\mathcal{E} := \{e_k : \{0, 1\}^k \to \{0, 1\}^n\}_{k \in \mathbb{N}_+}$ containing one code e_k for each possible message length k. The following result, due to Gilbert [35], can be used to construct an infinite sequence of codes with constant rate and distance.

Theorem 5 (*Gilbert-Varshamov Bound*) *For every* $k \in \mathbb{N}_+$, $0 \leq d^{\text{REL}} < \frac{1}{2}$ *and* $n \geq k / (1 - \mathcal{H}_2(d^{\text{REL}}))$, *there exists a code* $e : \{0, 1\}^k \to \{0, 1\}^n$ *with* $\text{dist}(e) = d^{\text{REL}}$, *where*

$$\mathcal{H}_2(d^{\text{REL}}) := d^{\text{REL}} \log_2 \left(\frac{1}{d^{\text{REL}}} \right) + (1 - d^{\text{REL}}) \log_2 \left(\frac{1}{1 - d^{\text{REL}}} \right).$$

Moreover, such a code can be computed in time $2^{O(n)}$.

5 Maximum e-Feasible Subsystem of Linear Inequalities

First, we prove the following auxiliary results that follow from Lemma 2 and will be useful in the remainder of the paper. Omitted proofs can be found in the Appendix.

Lemma 3 *Given a 3SAT formula φ,*

if φ is unsatisfiable, then for each (possibly randomized) Merlin$_2$'s strategy η_2 there exists a set S_i such that each Merlin$_1$'s assignment to variables in S_i is in conflict with Merlin$_2$'s assignment with a probability of at least $\rho/2v$.

Now, we introduce the *maximum ϵ -feasible subsystem of linear inequalities* problem. Given a system of linear inequalities $A\mathbf{x} \geq 0$ with $A \in [-1, 1]^{n_{\text{row}} \times n_{\text{col}}}$ and $\mathbf{x} \in \Delta_{n_{\text{col}}}$, we study the problem of finding the largest subsystem of linear inequalities that violate the constraints of at most ϵ. As we will show in Sect. 6, this problem presents some deep analogies with the problem of determining *good* posteriors in Bayesian persuasion problems.

Definition 3 (MFS) Given a matrix $A \in [-1, 1]^{n_{\text{row}} \times n_{\text{col}}}$, the problem of finding the *maximum feasible subsystem of linear inequalities* (MFS) reads as follows:

$$\max_{\mathbf{x}^* \in \Delta_{n_{\text{col}}}} \sum_{i \in [n_{\text{row}}]} I[w_i^* \geq 0] \quad \text{s.t. } \mathbf{w}^* = A\mathbf{x}^*.$$

We are interested in the problem of finding a vector **x** which yields at least the same number of feasible inequalities of MFS under a relaxation of the constraints with respect to Definition 3.

Definition 4 (ϵ-MFS) Given a matrix $A \in [-1, 1]^{n_{\text{row}} \times n_{\text{col}}}$, let

$$k^* := \max_{\mathbf{x}^* \in \Delta_{n_{\text{col}}}} \sum_{i \in [n_{\text{row}}]} I[w_i^* \geq 0] \text{ s.t. } \mathbf{w}^* = A \mathbf{x}^*.$$

Then, the problem of finding the *maximum ϵ-feasible subsystem of linear inequalities* (ϵ-MFS) amounts to finding a probability vector $\mathbf{x} \in \Delta_{n_{\text{col}}}$ such that, by letting $\mathbf{w} = A \mathbf{x}$, it holds:

$$\sum_{i \in [n_{\text{row}}]} I[w_i \geq -\epsilon] \geq k^*.$$

This problem was previously studied by Cheng et al. [5]. In particular, they design a PTAS for the ϵ-MFS problem guaranteeing the satisfaction of at least $k^* - \epsilon \cdot n_{\text{row}}$ inequalities. This yields a bi-criteria PTAS for the MFS problem.

5.1 Upper-bound on ϵ-MFS

First, we show that ϵ-MFS can be exactly solved in $n^{O(\log n)}$ steps for every fixed $\epsilon > 0$. We introduce the following auxiliary definition.

Definition 5 (k-uniform distribution) A probability distribution $\mathbf{x} \in \Delta_X$ is *k-uniform* if and only if it is the average of a multiset of k basis vectors in an $|X|$-dimensional space.

Equivalently, each entry x_i of a k-uniform distribution has to be a multiple of $1/k$. Then, the following result holds.

Proposition 6 ϵ-MFS *can be solved in $n^{O(\log n)}$ steps.*

Proof Denote by $\mathbf{x}^*$ the optimal solution of ϵ-MFS.

Let $\tilde{\mathbf{x}} \in \Delta_{n_{\text{col}}}$ be the empirical distribution of k i.i.d. samples drawn from probability distribution $\mathbf{x}^*$.

Moreover, let $\mathbf{w}^* := A \mathbf{x}^*$ and $\tilde{\mathbf{w}} := A \tilde{\mathbf{x}}$.

By Hoeffding's inequality we have

$$\Pr(w_i^* - \tilde{w}_i \geq \epsilon) \leq e^{-2k\epsilon^2}$$

for each $i \in [n_{\text{row}}]$.

Then, by the union bound, we get

$$\Pr(\exists i \text{ s.t. } w_i^* - \tilde{w}_i \geq \epsilon) \leq n_{\text{row}} \cdot e^{-2k\epsilon^2}.$$

Finally, we can write

$$\Pr(w_i^* - \tilde{w}_i \le \epsilon \ \forall i \in [n_{\text{row}}]) \ge 1 - n_{\text{row}} \cdot e^{-2k\epsilon^2}.$$

Thus, setting $k = \log n_{\text{row}}/\epsilon^2$ ensures the existence of a vector $\tilde{x}$ guaranteeing that, if $w_i^* \ge 0$, then $\tilde{w}_i \ge -\epsilon$.

Since $\tilde{x}$ is k-uniform by construction, we can find it by enumerating over all the $O((n_{\text{col}})^k)$ k-uniform probability vectors where $k = \log n_{\text{row}}/\epsilon^2$. Trivially, this task can be performed in $n^{\log n_{\text{row}}/\epsilon^2}$ steps and, therefore, in $n^{O(\log n)}$ steps. $\qquad\square$

5.2 Lower-bound on ϵ-MFS

Now we show that ϵ-MFS requires at least $n^{\tilde{\Omega}(\log n)}$ steps. In doing so, we close the gap with the upper bound stated by Proposition 6 except for polylogarithmic factors of $\log n$ in the denominator of the exponent.

Theorem 7 *Assuming ETH, there exists a constant $\epsilon > 0$ such that solving ϵ-MFS requires time $n^{\tilde{\Omega}(\log n)}$.*

Proof OVERVIEW. We provide a polynomial-time reduction from FREEGAME$_\delta$ (Definition 1) to ϵ-MFS, where $\epsilon = \delta/26 = \rho/(52\nu)$ (see Sect. 4.1 for the definition of parameters δ, ρ, ν). We show that, given a free game instance $\mathcal{F}_\varphi$, it is possible to build a matrix A such that, for a certain value k^*, the following holds:

(i) If $\omega(\mathcal{F}_\varphi) = 1$, then there exists a vector $\mathbf{x}$ such that

$$\sum_{i \in [n_{\text{row}}]} I[w_i \ge 0] = k^*, \tag{4}$$

where $\mathbf{w} = A\mathbf{x}$;

(ii) If $\omega(\mathcal{F}_\varphi) \le 1 - \delta$, then all vectors $\mathbf{x}$ are such that

$$\sum_{i \in [n_{\text{row}}]} I[w_i \ge -\epsilon] < k^*, \tag{5}$$

with $\mathbf{w} = A\mathbf{x}$.

CONSTRUCTION. In the free game $\mathcal{F}_\varphi$, Arthur sends a set of variables S_i to Merlin$_1$ and a set of clauses T_j to Merlin$_2$, where $i, j \in [m]$, $m = \sqrt{n}\,\text{polylog}(n)$. Then, Merlin$_1$'s (resp., Merlin$_2$'s) answer is denoted by $\xi_1 \in \Xi_1$ (resp., $\xi_2 \in \Xi_2$). The system of linear inequalities used in the reduction has a vector of variables $\mathbf{x}$ structured as follows.

1. *Variables corresponding to* Merlin$_2$'s *answers.* There is a variable x_{T_j,ξ_2} for each $j \in [m]$ and, due to Lemma 1 and assuming $|T_j| = 2m$, it holds $\xi_2 \in \Xi_2 = \{0, 1\}^{6m}$ (if $|T_j| < 2m$, we extend ξ_2 with a sufficient number of extra bits).

2. *Variables corresponding to* Merlin$_1$'s *answers.* We need to introduce some further machinery to augment the dimensionality of Ξ_1 through a viable mapping. Let $e : \{0, 1\}^{2m} \rightarrow \{0, 1\}^{8m}$ be the code defined in Theorem 5 with rate $1/4$ and relative distance $\mathrm{dist}(e) \geq 1/5$. We can safely assume that $|S_i| = 2m$ and $\xi_1 \in \Xi_1 = \{0, 1\}^{2m}$ (if $|S_i| < 2m$, we extend ξ_1 with a sufficient number of extra bits). Then, $e(\xi_1)$ is the $8m$-dimensional encoding of answer ξ_1 via code e. Let $e(\xi_1)_j$ be the j-th bit of vector $e(\xi_1)$. We have a variable $x_{i,\ell}$ for each index $i \in [8m]$ and $\ell := (\ell_j)_{j \in [m]} \in \{0, 1\}^m$. These $x_{i,\ell}$ variables can be interpreted as follows. Suppose to have an encoding of an answer for each of the possible set S_j. There are m such encodings, each of them having $8m$ bits. Then, it holds $x_{i,\ell} > 0$ if and only if the i-th bit of the encoding corresponding to S_j is ℓ_j.

There is a total of $m \cdot 2^m \cdot (2^{5m} + 8)$ variables. Matrix A has a number of columns equal to the number of variables. We denote with $A_{\cdot,(T_j,\xi_2)}$ the entry in row $(\cdot)$ and column corresponding to variable x_{T_j,ξ_2}. Analogously, $A_{\cdot,(i,\ell)}$ is the entry in row $(\cdot)$ and column corresponding to variable $x_{i,\ell}$. Rows are grouped in four types, denoted by $\{t_i\}_{i=1}^4$. We write $A_{t_i,\cdot}$ when referring to an entry of *any* row of type t_i. Further arguments may be added as a subscript to identify specific entries of A. Rows are structured as follows.

1. *Rows of type* t_1: there are q rows of type t_1 such that $A_{t_1,(T_j,\xi_2)} = 1$ for each $j \in [m], \xi_2 \in \Xi_2$, and -1 otherwise (the value of q is specified later in the proof).
2. *Rows of type* t_2: there are q rows for each subset $\mathcal{T} \subseteq \{T_j\}_{j \in [m]}$ with cardinality $m/2$ (i.e., there is a total of $q \cdot \binom{m}{m/2}$ rows of type t_2). Then, the following holds for each $\mathcal{T}$:

$$A_{(t_2,\mathcal{T}),(T_j,\xi_2)} = \begin{cases} -1 & \text{if } T_j \in \mathcal{T}, \xi_2 \in \Xi_2 \\ -1 & \text{if } T_j \notin \mathcal{T}, \xi_2 \in \Xi_2 \end{cases} \quad \text{and}$$

$$A_{(t_2,\mathcal{T}),(i,\ell)} = \quad 0 \quad \text{for each } i \in [8m], \ell \in \{0, 1\}^m.$$

3. *Rows of type* t_3: there are q rows of type t_3 for each subset of $4m$ indices $\mathcal{I}$ drawn from $[8m]$, for a total of $q \cdot \binom{8m}{4m}$ rows. For each subset of indices $\mathcal{I}$ we have:

$$A_{(t_3,\mathcal{I}),(T_j,\xi_2)} = 0 \quad \text{for each } T_j, \xi_2 \text{ and}$$

$$A_{(t_3,\mathcal{I}),(i,\ell)} = \begin{cases} -1 & \text{if } i \in \mathcal{I}, \ell \in \{0, 1\}^m \\ -1 & \text{if } i \notin \mathcal{I}, \ell \in \{0, 1\}^m \end{cases}.$$

4. *Rows of type* t_4: there is a row of type t_4 for each S_i and ξ_1. Each of these rows is such that:

$$A_{(t_4,S_i,\xi_1),(T_j,\xi_2)} = \begin{cases} -1/2 & \text{if } \mathcal{V}(S_i, T_j, \xi_1, \xi_2) = 1 \\ -1 & \text{otherwise} \end{cases} \quad \text{and}$$

$$A_{(t_4,S_i,\xi_1),(j,\ell)} = \begin{cases} -1/2 & \text{if } e(\xi_1)_j = \ell_i \\ -1 & \text{otherwise} \end{cases}.$$

Finally, we set $k^* = \left(1 + \binom{m}{m/2} + \binom{8m}{4m}\right) q + m$ and $q \gg m$ (for example, $q = 2^{10m}$). We say that row i satisfies ϵ-MFS condition for a certain $\mathbf{x}$ if $w_i \geq -\epsilon$, where $\mathbf{w} = A\mathbf{x}$ (in the following, we will also consider $w_i \geq 0$ as an alternative condition). We require at least k^* rows to satisfy the ϵ-MFS condition. Then, all rows of types t_1, t_2, t_3 and at least m rows of type t_4 must be such that w_i satisfies the ϵ-MFS condition.

COMPLETENESS. Given a satisfiable assignment of variables ζ to φ, we build vector $\mathbf{x}$ as follows. Let ζ_{T_j} be the partial assignment obtained by restricting ζ to the variables in the clauses of T_j (if $|T_j| < 2m$ we pad ζ_{T_j} with bits 0 until ζ_{T_j} has length $6m$). Then, we set $x_{T_j,\zeta_{T_j}} = 1/2m$. Moreover, for each $i \in [8m]$ and $\ell^i = (e(\zeta_{S_1})_i, \ldots, e(\zeta_{S_m})_i)$, we set $x_{i,\ell^i} = 1/16m$. We show that $\mathbf{x}$ is such that there are at least k^* rows i with $w_i \geq 0$ (Condition (4)). First, each row i of type t_1 is such that $w_i = 0$ since $\sum_{T_j,\xi_2} x_{T_j,\xi_2} = \sum_{i,\ell} x_{i,\ell} = 1/2$. For each T_j, $\sum_{\xi_2} x_{T_j,\xi_2} = 1/2\,m$. Then, for each subset $\mathcal{T}$ of $\{T_j\}_{j\in[m]}$, we have $\sum_{\xi_2, T_j \in \mathcal{T}} x_{T_j,\xi_2} = 1/4$. This implies that each row i of type t_2 is such that $w_i = 0$. A similar argument holds for rows of type t_3. Finally, we show that for each S_i there is at least a row i of type t_4 such that $w_i \geq 0$. Take the row corresponding to (S_i, ζ_{S_i}). For each $x_{b,\ell} > 0$ where $b \in [8\,m]$ and $\ell \in \{0,1\}^m$, it holds $e(\zeta_{S_i})_b = \ell_i$. Then, there are $8m$ columns played with probability $1/16\,m$ with value $1/2$, i.e., $\sum_{b,\ell} A_{(\mathsf{t}_4,S_i,\zeta_{S_i}),(b,\ell)} x_{b,\ell} = 1/4$. Moreover, for each (T_j, ζ_{T_j}), it holds $\mathcal{V}(S_i, T_j, \zeta_{S_i}, \zeta_{T_j}) = 1$. Then, $\sum_{T_j,\xi_2} A_{(\mathsf{t}_4,S_i,\zeta_{S_i}),(T_j,\zeta_{T_j})} x_{T_j,\xi_2} = -1/4$. This concludes the proof of completeness.

SOUNDNESS. We show that, if $\omega(\mathcal{F}_\varphi) \leq 1 - \delta$, there is no probability distribution $\mathbf{x}$ such that

$$\sum_{i \in n_{\mathrm{row}}} I[w_i \geq -\epsilon] \geq k^*, \tag{6}$$

with $\mathbf{w} = A\mathbf{x}$. Assume, by contradiction, that one such vector $\mathbf{x}$ exists. For the sake of clarity, we summarize the structure of the proof:

(i) We show that the probability assigned by $\mathbf{x}$ to columns with index (T_j, ξ_2) has to be *close* to $1/2$, and the same has to hold for columns of type (i, ℓ).

(ii) We show that $\mathbf{x}$ has to assign probability *almost* uniformly among T_js and indices i of the encoding of Ξ_1 (resp., Lemmas 5 and 6 below). Intuitively, this resembles the fact that, in $\mathcal{F}_\varphi$, Arthur draws questions T_j according to a uniform probability distribution.

(iii) For each S_i, there is at most one row $(\mathsf{t}_4, S_i, \xi_1)$ such that $w_{(\mathsf{t}_4,S_i,\xi_1)} \geq -\epsilon$ (Lemma 7). Together with the hypothesis that at least m rows of type t_4 satisfy the ϵ-MFS condition, this implies that there exists exactly one such row for each S_i.

(iv) Finally, we show that the above construction leads to a contradiction with Lemma 3 for a suitable free game.

Before providing the details of the four above steps, we introduce the following result, due to Babichenko et al. [34].

Lemma 4 *(Lemma 2 of Babichenko et al. [34]) Let* $\mathbf{v} \in \Delta^n$ *be a probability vector, and* $\mathbf{u}$ *be the n-dimensional uniform probability vector.*

If $||\mathbf{v} - \mathbf{u}|| > c$, *then there exists a subset of indices* $\mathcal{I} \subseteq [n]$ *such that* $|\mathcal{I}| = n/2$ *and* $\sum_{i \in \mathcal{I}} v_i > \frac{1}{2} + \frac{c}{4}$.

Then, we proceed with the following steps (the proofs of the auxiliary results can be found in Appendix A.2):

(i) Equation 6 requires all rows i of type t_1, t_2, t_3 to be such that $w_i \geq -\epsilon$. This implies that, for rows of type t_1, it holds

$$\sum_{T_j, \xi_2} x_{T_j, \xi_2} \geq \frac{1}{2}(1 - \epsilon). \tag{7}$$

If, by contradiction, this inequality did not hold, each row i of type t_1 would be such that $w_i < 1/2 - \epsilon/2 - (1/2 + \epsilon/2) = -\epsilon$, thus violating Eq. 6. Moreover, Eq. 6 implies that at least a row (t_4, S_i, ξ_1) has $w_{(t_4, S_i, \xi_1)} \geq -\epsilon$. Therefore, it holds $\sum_{i,\ell} x_{i,\ell} \geq 1/2 - \epsilon$. Indeed, if, by contradiction, this condition did not hold, all rows of type t_4 would have $w_i < 1/2(1/2 - \epsilon) - 1/2(1/2 + \epsilon) = -\epsilon$.

(ii) Let $\mathbf{v}_1 \in \Delta_m$ be the probability vector defined as

$$v_{1,j} := \frac{\sum_{\xi_2} x_{T_j, \xi_2}}{\sum_{j, \xi_2} x_{T_j, \xi_2}},$$

and $\tilde{\mathbf{v}}$ be a uniform probability vector of suitable dimension. The following result shows that having a bounded element-wise difference between $\mathbf{v}_1$ and $\tilde{\mathbf{v}}$ is a necessary condition for Eq. 6 to be satisfied.

Lemma 5 *If* $||\mathbf{v}_1 - \tilde{\mathbf{v}}||_1 > 16\epsilon$, *there exists a row* i *of type* t_2 *such that* $w_i < -\epsilon$.

Let $\mathbf{v}_2 \in \Delta_{[8m]}$ be the probability vector defined as

$$v_{2,i} := \frac{\sum_{\ell} x_{i,\ell}}{\sum_{i,\ell} x_{i,\ell}},$$

and $\tilde{\mathbf{v}}$ be a suitable uniform probability vector. Moreover, the following holds.

Lemma 6 *If* $||\mathbf{v}_2 - \tilde{\mathbf{v}}||_1 > 16\epsilon$, *there exists a row* i *of type* t_3 *such that* $w_i < -\epsilon$.

In order to satisfy Eq. 6, all rows i of type t_2 and t_3 have to be such that $w_i \geq -\epsilon$. Then, by Lemmas 5 and 6, it holds that $||\mathbf{v}_1 - \tilde{\mathbf{v}}||_1 \leq 16\epsilon$ and $||\mathbf{v}_2 - \tilde{\mathbf{v}}||_1 \leq 16\epsilon$.

(iii) We show that, for each S_i, there exists at most one row (t_4, S_i, ξ_1) for which $w_{(t_4, S_i, \xi_1)} \geq -\epsilon$.

Lemma 7 *For each* S_i, $i \in [m]$, *there exists at most one row* (t_4, S_i, ξ_1) *such that* $w_{(t_4, S_i, \xi_1)} \geq -\epsilon$.

Then, there are at least m rows (t_4, S_i, ξ_1) such that $w_{(\text{t}_4, S_i, \xi_1)} \geq -\epsilon$ and, by Lemma 7, we get that there exists exactly one such row for each S_i, $i \in [m]$. Therefore, for each S_i, there exists $\xi_1^i \in \Xi_1$ such that

$$\sum_{(T_j, \xi_2): \mathcal{V}(S_i, T_j, \xi_1^i, \xi_2)=1} x_{(T_j, \xi_2)} \geq \frac{1}{2} - 4\epsilon.$$

Notice that, if this condition did not hold, by Step (i) we would obtain

$$w_{\text{t}_4, S_i, \xi_1^i} < -\frac{1}{2}\left(\frac{1}{2} - 4\epsilon\right) - \frac{7}{2}\epsilon + \frac{1}{2}\left(\frac{1}{2} + \frac{\epsilon}{2}\right) = -\epsilon,$$

which would go against the satisfiability of Eq. 6.

(iv) Finally, let $\mathcal{F}_\varphi^*$ be a free game in which Arthur (i.e., the verifier) chooses question T_j with probability $v_{1,j}$ as defined in Step (ii), and Merlin$_2$ (i.e., the second prover) answers ξ_2 with probability $x_{T_j, \xi_2}/v_{1,j}$. In this setting (i.e., $\mathcal{F}_\varphi^*$), given question S_i to Merlin$_1$, the two provers will provide compatible answers with probability

$$\mathbb{P}\left(\mathcal{V}^*\left(S_i, T_j, \xi_1^i, \xi_2\right) = 1 \mid S_i\right) = \frac{1/2 - 4\epsilon}{\sum_{j, \xi_2} x_{T_j, \xi_2}} \geq \frac{1/2 - 4\epsilon}{1/2 + \epsilon} \geq 1 - 10\epsilon,$$

where the first inequality holds for Eq. 7 at Step (i). In a canonical (as in Definition 1) free game $\mathcal{F}_\varphi$, Arthur picks questions according to a uniform probability distribution. Therefore, the main difference between $\mathcal{F}_\varphi$ and $\mathcal{F}_\varphi^*$ is that, in the latter, Arthur draws questions for Merlin$_2$ from $\mathbf{v}_1$ which may not be a uniform probability distribution. However, we know that differences between $\mathbf{v}_1$ and a uniform probability vector must be limited. Specifically, by Lemma 5, we have $||\mathbf{v}_1 - \tilde{\mathbf{v}}||_1 \leq 16\epsilon$. Then, if Merlin$_1$ and Merlin$_2$ applied in $\mathcal{F}_\varphi$ the strategies we described for $\mathcal{F}_\varphi^*$, their answers would be compatible with probability at least $\mathbb{P}(\mathcal{V}(S_i, T_j, \xi_1^i, \xi_2) = 1 \mid S_i) \geq 1 - 26\epsilon$, for each S_i. Finally, by picking $\epsilon = \rho/52\,\nu$, we reach a contradiction with Lemma 3.

This concludes the proof. □

6 Hardness of (α, ϵ)-Persuasion

We show that a public signaling scheme approximating the value of the optimal one cannot be computed in polynomial time even if we allow it to be ϵ-persuasive (see Eq. 2). Specifically, assuming ETH, computing an (α, ϵ)-persuasive signaling scheme requires at least $n^{\tilde{\Omega}(\log n)}$, where the dimension of the instance is $n = O(\bar{n}\,d)$. We prove this result for the specific case of the k-voting problem, as introduced in Sect. 3. Besides its practical applicability, this problem is particularly instructive in highlighting the strong connection between the problem of finding suitable posteriors and the ϵ-MFS problem, as discussed in the following lemma. An analogous observation was also highlighted by Cheng et al. in [5].

Lemma 8 *Given a k-voting instance, the problem of finding a posterior* $\mathbf{p} \in \Delta_\Theta$ *such that* $W_\epsilon(\mathbf{p}) \geq k$ *is equivalent to finding an ϵ-feasible subsystem of k linear inequalities over the simplex when* $A \in [-1, 1]^{\bar{n} \times d}$ *is such that:*

$$A_{r,\theta} = u_\theta^r(a_0) - u_\theta^r(a_1) \quad \text{for each } r \in \mathcal{R}, \theta \in \Theta. \tag{8}$$

Proof By setting $\mathbf{x} = \mathbf{p}$, it directly follows that $\sum_{i \in [\bar{n}]} I[A_i \mathbf{x} \geq -\epsilon] \geq k$ if and only if $W_\epsilon(\mathbf{p}) \geq k$. □

The above lemma shows that deciding if there exists a posterior $\mathbf{p}$ such that $W(\mathbf{p}) \geq k$ or if *all* the posteriors have $W_\epsilon(\mathbf{p}) < k$ (i.e., deciding if the utility of the sender can be greater than zero) is as hard as solving the ϵ-MFS problem. More precisely, if an ϵ-MFS instance does not admit any solution, then there does not exist a posterior guaranteeing a strictly positive winning probability for the sender's preferred candidate. On the other hand, if an ϵ-MFS instance admits a solution, there exists a signaling scheme where at least one of the induced posteriors guarantees strictly positive winning probability to the sender's preferred candidate. However, the above connection between the ϵ-MFS problem and the k-voting problem is not sufficient to prove the inapproximability of the k-voting problem, as the probability whereby this posterior is reached may be arbitrarily small.

Luckily enough, the next theorem shows that it is possible to strengthen the inapproximability result by constructing an instance in which, when 3SAT is satisfiable, there exists a signaling scheme such that all the induced posteriors satisfy $W(\mathbf{p}) \geq k$ (i.e., the sender's preferred candidate wins with a probability of 1). The main idea is to suitably extend the construction of Theorem 7 with an additional set of states $\{\theta_\mathbf{d}\}_\mathbf{d} \in \{0, 1\}^{7m}$, where we can see vectors $\mathbf{d}$ as the concatenation of a subvector $\mathbf{d}_S \in \{0, 1\}^m$ and a subvector $\mathbf{d}_T \in \{0, 1\}^{6m}$. Moreover, we need to extend the set of receivers. In particular, we replace each receiver relative to a set S_i and an answer ξ_1 with a set including a receiver for each $\mathbf{d}$. The new receivers' payoffs are defined as follows. Let $\oplus$ be the XOR operator. The payoff of the receiver relative to S_i, ξ_1, and $\mathbf{d}$ in a state $\theta_{(T_j, \xi_2 \oplus d_T)}$ is equivalent to the payoff of the original receiver in the state $\theta_{(T_j, \xi_2)}$, while we use a similar procedure for the payoffs in the states $\theta_{(i, \ell)}$. Then, the signaling scheme employs a signal $s_\mathbf{d}$ for each $\mathbf{d}$. Each signal $s_\mathbf{d}$ defines which of the $\{0, 1\}^{7m}$ games we are playing. All these games are equivalent since we are simply changing the meaning of the states: for example, a state $\theta_{(T_j, \xi_2 \oplus d_T)}$ is equivalent to the original state $\theta_{(T_j, \xi_2)}$. Using this construction, we have that all the signals induce a posterior in which at least k voters votes for c_0, while in the original game only one signal induces a posterior that satisfies this condition.

Theorem 8 *Given a k-voting instance and assuming ETH, there exists a constant* $\epsilon^* > 0$ *such that, for any* $0 < \epsilon \leq \epsilon^*$, *finding an* (α, ϵ)-*persuasive signaling scheme requires* $n^{\tilde{\Omega}(\log n)}$ *steps for any multiplicative or additive factor* $\alpha \in (0, 1)$.

Proof OVERVIEW. By following the proof of Theorem 7, we can provide a polynomial-time reduction from FREEGAME$_\delta$ to the problem of finding an ϵ-persuasive signaling scheme in k-voting, with $\epsilon = \delta/780 = \rho/1560\nu$. Specifically, if $\omega(\mathcal{F}_\varphi) = 1$, there exists a signaling scheme guaranteeing the sender an expected value of 1. Otherwise,

if $\omega(\mathcal{F}_\varphi) \leq 1 - \delta$, then all posteriors are such that $W_\epsilon(\mathbf{p}) < k$ (i.e., the sender cannot obtain more than 0).

CONSTRUCTION. The k-voting instance has the following possible states of nature.

1. $\theta_{(T_j,\xi_2)}$ for each set of clauses T_j, $j \in [m]$, and answer $\xi_2 \in \Xi_2 = \{0, 1\}^{6m}$.
   Let $e : \{0, 1\}^{2m} \to \{0, 1\}^{8m}$ be an encoding function with $R = 1/4$ and dist$(e) \geq 1/5$ (as in the proof of Theorem 7).
   We have a state $\theta_{(i,\ell)}$ for each $i \in [8m]$, and $\ell = (\ell_1, \ldots, \ell_m) \in \{0, 1\}^m$.
2. There is a state $\theta_{\mathbf{d}}$ for each $\mathbf{d} \in \{0, 1\}^{7m}$.
   It is useful to see vector $\mathbf{d}$ as the union of the subvector $\mathbf{d}_S \in \{0, 1\}^m$ and the subvector $\mathbf{d}_T \in \{0, 1\}^{6m}$.

The common prior μ is such that:

$$\mu_{\theta_{(T_j,\xi_2)}} = \frac{1}{m\, 2^{2+6m}} \quad \text{for each } \theta_{(T_j,\xi_2)},$$

$$\mu_{\theta_{(i,\ell)}} = \frac{1}{m\, 2^{5+m}} \quad \text{for each } \theta_{(i,\ell)},$$

$$\mu_{\theta_{\mathbf{d}}} = \frac{1}{2^{1+7m}} \quad \text{for each } \theta_{\mathbf{d}}.$$

To simplify the notation, in the remaining of the proof, let $u_\theta^r := u_\theta^r(a_0) - u_\theta^r(a_1)$. The k-voting instance comprises the following receivers.

1. *Receivers of type* t_1: there are q (the value of q is specified later in the proof) receivers of type t_1, which are such that $u_{\theta_{(T_j,\xi_2)}}^{\mathsf{t}_1} = 1$ for each (T_j, ξ_2), and $-1/3$ otherwise.
2. *Receivers of type* t_2: there are q receivers of type t_2 such that $u_{\theta_{(i,\ell)}}^{\mathsf{t}_2} = 1$ for each (i, ℓ), and $-1/3$ otherwise.
3. *Receivers of type* t_3: there are q receivers of type t_3 for each subset $\mathcal{T} \subseteq \{T_j\}_{j\in[m]}$ of cardinality $m/2$. Each receiver corresponding to the subset $\mathcal{T}$ is such that:

$$u_{\theta_{(T_j,\xi_2)}}^{(\mathsf{t}_3,\mathcal{T})} = \begin{cases} -1 & \text{if } T_j \in \mathcal{T}, \xi_2 \in \Xi_2 \\ -1 & \text{if } T_j \notin \mathcal{T}, \xi_2 \in \Xi_2 \end{cases} \quad \text{and } u_\theta^{(\mathsf{t}_3,\mathcal{T})} = 0 \text{ for every other} \theta.$$

4. *Receivers of type* t_4:
   we have q receivers of type t_4 for each subset $\mathcal{I}$ of $4m$ indices selected from $[8m]$. Each receiver corresponding to subset $\mathcal{I}$ is such that:

$$u_{\theta_{(i,\ell)}}^{(\mathsf{t}_4,\mathcal{I})} = \begin{cases} -1 & \text{if } i \in \mathcal{I}, \ell \in \{0, 1\}^m \\ -1 & \text{if } i \notin \mathcal{I}, \ell \in \{0, 1\}^m \end{cases} \quad \text{and } u_\theta^{(\mathsf{t}_4,\mathcal{I})} = 0 \text{ for every other} \theta.$$

5. *Receivers of type* t_5:
   there is a receiver of type t_5 for each S_i, $\xi_1 \in \Xi_1$ and $\mathbf{d} \in \{0, 1\}^{7m}$.

Then, for each receiver of type t_5 the following holds:

$$u_\theta^{(t_5, S_i, \xi_1, \mathbf{d})} = \begin{cases} -1/2 & \text{if } \theta = \theta_{(T_j, \xi_2)} \text{ and } \mathcal{V}(S_i, T_j, \xi_1, \xi_2 \oplus \mathbf{d}_T) = 1 \\ -1/2 & \text{if } \theta = \theta_{(i', \ell)} \text{ and } e(\xi_1)_{i'} = [\ell \oplus \mathbf{d}_S]_i \\ -1/2 & \text{if } \theta = \theta_{\mathbf{d}} \\ -1 & \text{otherwise} \end{cases}$$

Finally, we set $k = \left(2 + \binom{m}{m/2} + \binom{8m}{4m} \right) q + m$. By setting $q \gg m$ (e.g., $q = 2^{10m}$), candidate a_0 can get at least k votes only if all receivers of type t_1, t_2, t_3, t_4 vote for her.

COMPLETENESS. Given a satisfiable assignment ζ to the variables in φ, let $[\zeta]_{T_j} \in \{0, 1\}^{6m}$ be the vector specifying the variables assignment of each clause in T_j, and $[\zeta]_{S_i} \in \{0, 1\}^{2m}$ be the vector specifying the assignment of each variable belonging to S_i. The sender has a signal for each $\mathbf{d} \in \{0, 1\}^{7m}$. The set of signals is denoted by $\mathcal{S}$, where $|\mathcal{S}| = 2^{7m}$, and a signal is denoted by $s_\mathbf{d} \in \mathcal{S}$. We define a signaling scheme ϕ as follows. First, we set $\phi_{\theta_\mathbf{d}}(s_\mathbf{d}) = 1$ for each $\theta_\mathbf{d}$. If $|T_j| < 2m$ for some $j \in [m]$, we pad $[\zeta]_{T_j}$ with bits 0 util $|[\zeta]_{T_j}| = 6m$. Then, for each T_j, $\phi_{\theta_{(T_j, [\zeta]_{T_j} \oplus \mathbf{d}_T)}}(s_\mathbf{d}) = 1/2^m$. For each $i \in [8m]$, set $\phi_{\theta_{(i, \ell \oplus \mathbf{d}_S)}} = 1/2^{6m}$, where $\ell = (e([\zeta]_{S_1})_i, \ldots, e([\zeta]_{S_m})_i)$. First, we prove that the signaling scheme is well-formed. For each state $\theta_{(T_j, \xi_2)}$, it holds that

$$\sum_{s_\mathbf{d} \in \mathcal{S}} \phi_{\theta_{(T_j, \xi_2)}}(s_\mathbf{d}) = \frac{1}{2^m} \left| \{ \mathbf{d} : [\zeta]_{T_j} \oplus \mathbf{d}_T = \xi_2 \} \right| = 1,$$

and, for each $\theta_{(i, \ell)}$, the following holds:

$$\sum_{s_\mathbf{d} \in \mathcal{S}} \phi_{\theta_{(i, \ell)}}(s_\mathbf{d}) = \frac{1}{2^{6m}} \left| \{ \mathbf{d} : (e([\zeta]_{S_1})_i, \ldots, e([\zeta]_{S_m})_i \oplus \mathbf{d}_S = \ell \} \right| = 1.$$

Now, we show that there exist at least k voters that will choose a_0. Let $\mathbf{p} \in \Delta_\Theta$ be the posterior induced by a signal $s_\mathbf{d}$. All receivers of type t_1 choose a_0 since it holds:

$$\sum_{(T_j, \xi_2)} p_{\theta_{(T_j, \xi_2)}} = \frac{\sum_{(T_j, \xi_2)} \mu_{\theta_{(T_j, \xi_2)}} \phi_{\theta_{(T_j, \xi_2)}}(s_\mathbf{d})}{\sum_{\theta \in \Theta} \mu_\theta \phi_\theta(s_\mathbf{d})}$$

$$= \frac{1}{2^{2+7m}} \left(\frac{1}{2^{1+7m}} + \frac{1}{2^{2+7m}} + \frac{1}{2^{2+7m}} \right)^{-1}$$

$$= \frac{1}{4}.$$

Analogously, all receivers of type t_2 select a_0. Furthermore, for each T_j, it holds $\sum_{\xi_2} p_{\theta_{(T_j, \xi_2)}} = 1/4m$. Then, for each subset $\mathcal{T} \subseteq \{T_j\}_{j \in [m]}$ of cardinality $m/2$, it holds $\sum_{T_j \in \mathcal{T}, \xi_2} p_{\theta_{(T_j, \xi_2)}} = m/2 \; 1/4 \; m = 1/8$. Therefore, each receiver of type

t_3 chooses a_0. An analogous argument holds for receivers of type t_4. Finally, we show that, for each S_i, the receiver $(t_5, S_i, [\zeta]_{S_i}, \mathbf{d})$ chooses a_0. In particular, receiver $(t_5, S_i, [\zeta]_{S_i}, \mathbf{d})$ has the following expected utility:

$$\frac{1}{2} p_{\theta_\mathbf{d}} - \frac{1}{2} \sum_{(T_j, \xi_2)} p_{\theta_{(T_j, \xi_2)}} - \frac{1}{2} \sum_{(i', \ell)} p_{\theta_{(i', \ell)}} = 0$$

since, for each $p_{(T_j, \xi_2)} > 0$, the following holds $\xi_2 \oplus \mathbf{d}_T = [\zeta]_{T_j} \oplus \mathbf{d}_T \oplus \mathbf{d}_T = [\zeta]_{T_j}$ and $\mathcal{V}(S_i, T_j, [\zeta]_{S_i}, \xi_2 \oplus \mathbf{d}_T) = \mathcal{V}(S_i, T_j, [\zeta]_{S_i}, [\zeta]_{T_j}) = 1$ for each T_j. Moreover, for each $p_{(\theta_{i'}, l)} > 0$, it holds $[l \oplus d_S]_i = e([\zeta]_{S_i})_{i'} \oplus d_{S,i} \oplus d_{S,i} = e([\zeta]_{S_i})_{i'}$. This concludes the proof of completeness.[6]

SOUNDNESS. We prove that, if $\omega(\mathcal{F}_\varphi) \leq 1 - \delta$, there is no posterior in which a_0 is chosen by at least k receivers, thus implying that the sender's utility is equal to 0. Now, suppose, towards a contradiction, that there exists a posterior $\mathbf{p}$ such that at least k receivers select a_0. Let $\gamma := \sum_{(T_j, \xi_2)} p_{\theta_{(T_j, \xi_2)}} + \sum_{(i, \ell)} p_{\theta_{(i, \ell)}}$. Since all voters of types t_1 and t_2 vote for a_0, it holds that $\sum_{(T_j, \xi_2)} p_{\theta_{(T_j, \xi_2)}} \geq \frac{1}{4} - \epsilon$ and $\sum_{(i, \ell)} p_{\theta_{(i, \ell)}} \geq \frac{1}{4} - \epsilon$. Moreover, since at least a receiver $(t_5, S_i, \xi_1, \mathbf{d})$ must play a_0, there exists a $\mathbf{d} \in \{0, 1\}^{7m}$ and a state $\theta_\mathbf{d}$ with $p_\theta \geq \frac{1}{2} - \epsilon$. This implies that $\frac{1}{2} - 2\epsilon \leq \gamma \leq \frac{1}{2} + \epsilon$.

Consider the reduction to ϵ'-MFS, with $\epsilon' = \rho/52\nu$ (Theorem 7). Let $x_{(T_j, \xi_2)} = p_{\theta_{(T_j, \xi_2 \oplus \mathbf{d}_T)}}/\gamma$, $x_{(i, \ell)} = p_{\theta_{(i, \ell \oplus \mathbf{d}_S)}}/\gamma$, and $\epsilon = \epsilon'/30$. All rows of type t_1 of ϵ'-MFS are such that

$$w_{t_1} = \frac{1}{\gamma} \left(\sum_{(T_j, \xi_2)} p_{\theta_{(T_j, \xi_2)}} - \sum_{(i, \ell)} p_{\theta_{(i, \ell)}} \right) \geq -\frac{3\epsilon}{\gamma} \geq -9\epsilon \geq -\epsilon'.$$

All voters of type t_3 choose a_0. Then, for all $\mathcal{T} \subseteq \{T_j\}_{j \in [m]}$ of cardinality $m/2$, it holds:

$$\sum_{(T_j, \xi_2): T_j \in \mathcal{T}} p_{\theta_{(T_j, \xi_2)}} - \sum_{(T_j, \xi_2): T_j \notin \mathcal{T}} p_{\theta_{(T_j, \xi_2)}} \geq -\epsilon.$$

Then, all rows of type t_2 of ϵ'-MFS are such that:

$$w_{(t_2, \mathcal{T})} = \frac{1}{\gamma} \left(\sum_{(T_j, \xi_2): T_j \in \mathcal{T}} p_{\theta_{(T_j, \xi_2)}} - \sum_{(T_j, \xi_2): T_j \notin \mathcal{T}} p_{\theta_{(T_j, \xi_2)}} \right) \geq -\frac{\epsilon}{\gamma} \geq -3\epsilon \geq -\epsilon'.$$

A similar argument proves that all rows of type t_3 of the instance of ϵ'-MFS have $w_{(t_3, \mathcal{I})} \geq -\epsilon'$.

---

[6] For the sake of presentation, in the proof, we employ indirect signals of type $s_\mathbf{d}$. However, it is possible to construct an equivalent direct signaling scheme. Let $\mathbf{p}^\mathbf{d} \in \Delta_\Theta$ be the posterior induced by $s_\mathbf{d}$. Then, it is enough to substitute each $s_\mathbf{d}$ with a direct signal recommending a_0 to all receivers such that $\sum_\theta p_\theta^\mathbf{d} u_\theta^r \geq 0$, and a_1 to all the others.

To conclude the proof, we prove that, for each voter $(\mathtt{t}_5, S_i, \xi_1, \mathbf{d})$ that votes for a_0, the corresponding row $(\mathtt{t}_4, S_i, \xi_1)$ of the instance ϵ'-MFS is such that $w_{(\mathtt{t}_4, S_i, \xi_1)} \geq -\epsilon'$. Let $\gamma' := \sum_{(T_j, \xi_2): \mathcal{V}(S_i, T_j, \xi_1, \xi_2) = 1} x_{(T_j, \xi_2)}$ and $\gamma'' := \sum_{(i', \ell): e(\xi_1)_{i'} = \ell_i} x_{(i', \ell)}$. First, we have that $\gamma' \geq 1/4 - 7\epsilon$. If this did not hold, we would have

$$\sum_\theta p_\theta u_\theta^{(\mathtt{t}_5, S_i, \xi_1, \mathbf{d})} < -\frac{1}{2}\left(\frac{1}{4} - \epsilon\right) - \frac{1}{2}\left(\frac{1}{4} - 7\epsilon\right) - 6\epsilon + \frac{1}{2}\left(\frac{1}{2} + 2\epsilon\right) = \epsilon.$$

Similarly, it holds $\gamma'' \geq 1/4 - 7\epsilon$. Hence

$$
\begin{aligned}
w_{(\mathtt{t}_4, S_i, \xi_1)} &= -\frac{1}{2}\gamma' + \frac{1}{2}\gamma'' - (1 - \gamma' - \gamma'') \\
&= \frac{1}{2\gamma}\left(\sum_{(T_j, \xi_2): \mathcal{V}(S_i, T_j, \xi_1, \xi_2) = 1} p_{\theta_{(T_j, \xi_2 \oplus \mathbf{d}_T)}} + 3 \sum_{(i', \ell): e(\xi_1)_{i'} = \ell_i} p_{\theta_{(i', \ell \oplus \mathbf{d}_S)}}\right) - 1 \\
&\geq \frac{2(1/4 - 7\epsilon)}{1/2 + \epsilon} - 1 \geq -30\epsilon = -\epsilon'.
\end{aligned}
$$

Thus, there exists a probability vector $\mathbf{x}$ for the instance of ϵ'-MFS in which at least k rows satisfy the ϵ'-MFS condition (Eq. 5), which is in contradiction with $\omega(\mathcal{F}_\varphi) \leq 1 - \delta$. This concludes the proof. $\qquad\square$

Theorem 8 shows that, assuming the ETH, computing an (α, ϵ)-persuasive signaling schemes requires at least a quasi-polynomial number of steps in the specific scenario of a k-voting instance. Therefore, the same holds in the general setting of arbitrary public persuasion problems with binary action spaces, which is precisely the claim of Theorem 1.

7 A Quasi-Polynomial Time Algorithm

In this section, we prove that our hardness result (Theorem 8) is tight by devising a bi-criteria approximation algorithm. Our result extends the results by Cheng et al. [5] and Xu [22], which deal with signaling problems with binary action spaces and sender's utility functions which are independent from the state of nature. This is arguably a restrictive assumption, and even the original Bayesian persuasion framework by Kamenica and Gentzkow [14] describes state-dependent sender's utility functions. Our results generalize those by Cheng et al. [5] to the case of state-dependent sender's utility functions, and arbitrary discrete action spaces.

In order to prove our result, we need some further machinery. Let $\mathcal{Z}^r := 2^{\mathcal{A}^r}$ be the power set of $\mathcal{A}^r$. Then, $\mathcal{Z} := \times_{r \in \mathcal{R}} \mathcal{Z}^r$ is the set of tuples specifying a subset of $\mathcal{A}^r$ for each receiver r. For a given probability distribution over the states of nature, we are interested in determining the set of best responses of each receiver r, i.e., the subset of $\mathcal{A}^r$ maximizing her expected utility. Formally, we have the following.

Definition 6 (BR-set) Given a probability distribution over states of nature $\mathbf{p} \in \Delta_\Theta$, the *best-response set* (BR-set) $\mathcal{M}_\mathbf{p} := (Z^1, \ldots, Z^n) \in \mathcal{Z}$ is such that

$$Z^r = \arg\max_{a \in \mathcal{A}^r} \sum_{\theta \in \Theta} p_\theta \, u_\theta^r(a) \qquad \text{for each } r \in \mathcal{R}.$$

Similarly, we define a notion of ϵ-BR-set which comprises ϵ-approximate best responses to a given distribution over the states of nature.

Definition 7 (ϵ-BR-set) Given a probability distribution over states of nature $\mathbf{p} \in \Delta_\Theta$, the ϵ-*best-response set* (ϵ-BR-set) $\mathcal{M}_{\mathbf{p},\epsilon} := (Z^1, \ldots, Z^n) \in \mathcal{Z}$ is such that, for each $r \in \mathcal{R}$, action a belongs to Z^r if and only if

$$\sum_{\theta \in \Theta} p_\theta \, u_\theta^r(a) \geq \sum_{\theta \in \Theta} p_\theta \, u_\theta^r(a') - \epsilon \qquad \text{for each } a' \in \mathcal{A}^r.$$

We introduce a suitable notion of *approximability* of the sender's objective function. Our notion of α-*approximable function* is a generalization of the one proposed by Xu [22, Definition 4.5] to the setting of arbitrary action spaces and state-dependent sender's utility functions.

Definition 8 (α-Approximability) Let $f := \{f_\theta\}_{\theta \in \Theta}$ be a set of functions $f_\theta : \mathcal{A} \to [0, 1]$.

We say that f is α-approximable if there exists a function $g : \Delta_\Theta \times \mathcal{Z} \to \mathcal{A}$ computable in polynomial time such that, for all $\mathbf{p} \in \Delta_\Theta$ and $Z \in \mathcal{Z}$, it holds: $\mathbf{a} = g(\mathbf{p}, Z), \mathbf{a} \in Z$ and

$$\sum_{\theta \in \Theta} p_\theta \, f_\theta(\mathbf{a}) \geq \alpha \max_{\mathbf{a}^* \in Z} \sum_{\theta \in \Theta} p_\theta \, f_\theta(\mathbf{a}^*).$$

The voting function f defined in Sect. 3 is 1-approximable, while, for example, when the action space is binary a non-monotone submodular function is $1/2$-approximable. The α-approximability assumption is a natural requirement since, otherwise, even evaluating the sender's objective value would result in an intractable problem. When f is α-approximable, it is possible to find an approximation of the optimal receivers' actions profile when they are constrained to select actions profiles in Z.

We now provide an algorithm which computes in quasi-polynomial time, for any α-approximable f, a bi-criteria approximation of the optimal solution with an approximation on the objective value arbitrarily close to α. When f is 1-approximable our result yields a bi-criteria QPTAS for the problem. The key idea is showing that an optimal signaling scheme can be approximated by a convex combination of suitable k-uniform posteriors. As in previous works [5, 22], the key part of the proof is a decomposition lemma that proves that all the posteriors can be decomposed in a convex combination of k-uniform posteriors with a small loss in utility. However, the assumption of state-dependent sender's utility functions makes previous approaches ineffective in our setting. In particular, we observe that previous decomposition lemmas are based on a direct application of the Hoeffding's and union bounds. In our case,

such a direct derivation is not possible, and we need to introduce some technical inter-
mediate results (Lemmas 9 – 12). In particular, we need to develop a new probabilistic
analysis of the decomposition lemma. Let $\varrho := \max_{r \in \mathcal{R}} \varrho_r$, $\bar{n} := |\mathcal{R}|$, and $d := |\Theta|$.
The proof of our main positive result, as stated in Theorem 2, goes as follows.

Proof of Theorem 2 We show that there exists a poly $\left(d^{\log(\bar{n}\varrho/\delta)/\epsilon^2}\right)$ algorithm that
computes the given approximation. Let $k = 32 \log(4\bar{n}\varrho/\delta)/\epsilon^2$ and $\mathcal{K} \subset \Delta_\Theta$ be the
set of k-uniform distributions over Θ (Def. 5). We prove that all posteriors $\mathbf{p}^* \in \Delta_\Theta$ can
be decomposed as a convex combination of k-uniform posteriors without lowering too
much the sender's expected utility. Formally, each posterior $\mathbf{p}^* \in \Delta_\Theta$ can be written
as $\mathbf{p}^* = \sum_{\mathbf{p} \in \mathcal{K}} \gamma_{\mathbf{p}} \, \mathbf{p}$, with $\gamma \in \Delta_\mathcal{K}$ such that

$$\sum_{\mathbf{p} \in \mathcal{K}} \gamma_{\mathbf{p}} \sum_{\theta \in \Theta} p_\theta \, f_\theta(g(p, \mathcal{M}_\epsilon(p))) \geq \alpha \, (1 - \delta) \max_{\mathbf{a}^* \in \mathcal{M}(\mathbf{p}^*)} \sum_{\theta \in \Theta} p_\theta^* \, f_\theta(\mathbf{a}^*).$$

Let $\tilde{\gamma} \in \mathcal{K}$ be the empirical distribution of k i.i.d. samples from $\mathbf{p}^*$, where each
θ has probability p_θ^* of being sampled. Therefore, the vector $\tilde{\gamma}$ is a random variable
supported on k-uniform posteriors with expectation $\mathbf{p}^*$. Moreover, let $\gamma \in \Delta_\mathcal{K}$ be a
probability distribution such as, for each $\mathbf{p} \in \mathcal{K}$, $\gamma_{\mathbf{p}} := \Pr(\tilde{\gamma} = \mathbf{p})$. For each $\gamma \in \Delta_\mathcal{K}$
and $\mathbf{p} \in \mathcal{K}$, we denote by $\gamma_{\mathbf{p}}^{(\theta,i)}$ the conditional probability of having observed posterior
$\mathbf{p}$, given that the posterior must assign probability i/k to state θ. Formally, for each
$\mathbf{p} \in \mathcal{K}$, if $p_\theta = i/k$, we have

$$\gamma_{\mathbf{p}}^{(\theta,i)} = \frac{\gamma_{\mathbf{p}}}{\displaystyle\sum_{\mathbf{p}':p_\theta'=i/k} \gamma_{\mathbf{p}'}},$$

and $\gamma_{\mathbf{p}}^{(\theta,i)} = 0$ otherwise. The random variable $\tilde{\gamma}^{(\theta,i)} \in \mathcal{K}$ is such that, for each $\mathbf{p} \in \mathcal{K}$,
$\Pr(\tilde{\gamma}^{(\theta,i)} = \mathbf{p}) = \gamma_{\mathbf{p}}^{(\theta,i)}$. Finally, let $\mathcal{P} \subseteq \mathcal{K}$ be the set of posteriors such that

$$\mathcal{P} := \left\{ \mathbf{p} \in \mathcal{K} : \left| \sum_\theta p_\theta u_\theta^r(a) - \sum_\theta p_\theta^* u_\theta^r(a) \right| \leq \frac{\epsilon}{2} \; \forall r \in \mathcal{R}, a \in \mathcal{A}^r \right\}. \tag{9}$$

Now, we prove the following intermediate results (the proofs of the auxiliary results
are provided in Appendix A.3). The following lemma show that, given a posterior $\mathbf{p}^*$
and a state θ, if we take k i.i.d. samples from $\mathbf{p}^*$ and we consider only the induced
posteriors $\mathbf{p}$ in which p_θ is close to p_θ^*, then the probability that the utility of all the
receivers in $\mathbf{p}$ is close to the their utility in $\mathbf{p}^*$ is close to 1.

Lemma 9 *Given $\mathbf{p}^* \in \Delta_\Theta$, for each $\theta \in \Theta$, and for each $i \in [k]$ such that*

$$\left| \frac{i}{k} - p_\theta^* \right| \leq \frac{\epsilon}{4},$$

it holds:

$$\sum_{\mathbf{p}\in\mathcal{P}:p_\theta=i/k} \gamma_\mathbf{p} \geq \left(1 - \frac{\delta}{2}\right) \sum_{\mathbf{p}\in\mathcal{K}:p_\theta=i/k} \gamma_\mathbf{p},$$

where γ is the distribution of k i.i.d. samples from $\mathbf{p}^*$.

Then, we show that the condition in the previous lemma is satisfied with high probability. In particular, we show that given a posterior $\mathbf{p}^*$ and a state θ, if we take k i.i.d. samples from $\mathbf{p}^*$, then with probability close to 1 the induced posterior $\mathbf{p}$ is such that p_θ is close to p_θ^*. Formally, we prove the following lemma.

Lemma 10 *Given $\mathbf{p}^* \in \Delta_\Theta$, for each $\theta \in \Theta$, it holds:*

$$\sum_{i:|i/k-\mathbf{p}_\theta^*|\geq\epsilon/4} \quad \sum_{\mathbf{p}\in\mathcal{K}:p_\theta=i/k} \gamma_p \leq \frac{\delta}{2} p_\theta^*,$$

where γ is the distribution of k i.i.d. samples from $\mathbf{p}^*$.

The following result combines Lemmas 9 and 10. In particular, if we consider the distribution of k i.i.d. samples from a posterior $\mathbf{p}^*$, we have that, for each θ, the probability that in state θ the utility of all the receivers is close to their utility in $\mathbf{p}^*$ is close to 1. Equivalently, the induced posterior belongs to $\mathcal{P}$ as defined in (9).

Lemma 11 *Given a $\mathbf{p}^* \in \Delta_\Theta$, for each $\theta \in \Theta$, it holds:*

$$\sum_{\mathbf{p}\in\mathcal{P}} \gamma_p \, p_\theta \geq (1 - \delta) \, p_\theta^*,$$

where γ is the distribution of k i.i.d. samples from $\mathbf{p}^*$.

Now, we need to prove that all the posteriors in $\mathcal{P}$ guarantee to the sender at least the same expected utility of $\mathbf{p}^*$. Formally, we prove that the ϵ-BR-set of each $\mathbf{p} \in \mathcal{P}$ contains the BR-set of $\mathbf{p}^*$. This is shown via the following lemma.

Lemma 12 *Given $\mathbf{p}^* \in \Delta_\Theta$, for each $\mathbf{p} \in \mathcal{P}$, it holds: $\mathcal{M}(\mathbf{p}^*) \subseteq \mathcal{M}_\epsilon(\mathbf{p})$.*

Finally, we prove that we can represent each posterior $\mathbf{p}^*$ as a convex combination of k-uniform posteriors with a small loss in the sender's expected utility. For $\mathbf{p} \in \mathcal{K}$ and $Z \in \mathcal{Z}$, let $g^* : \Delta_\Theta \times \mathcal{Z} \to [0, 1]$ be a function such that

$$g^*(\mathbf{p}, Z) := \max_{\mathbf{a}\in Z} \sum_\theta p_\theta \, f_\theta(\mathbf{a}).$$

Given $\mathbf{p}^* \in \Delta_\Theta$, we are interested in bounding the difference in the sender's expected utility when $\mathbf{p}^*$ is approximated as a convex combination γ of k-uniform posteriors, the sender exploits an α-approximation of f, and receivers play ϵ-persuasive best-responses. Formally,

Lemma 13 *Given a* $\mathbf{p}^* \in \Delta_\Theta$, *it holds:*

$$\sum_{\mathbf{p} \in \mathcal{K}} \gamma_{\mathbf{p}} \sum_\theta p_\theta \, f_\theta(g(\mathbf{p}, \mathcal{M}_\epsilon(\mathbf{p}))) \geq \alpha(1 - \delta) f_\theta(g^*(\mathbf{p}^*, \mathcal{M}(\mathbf{p}^*))),$$

where γ *is the distribution of* k *i.i.d. samples from* $\mathbf{p}^*$.

Therefore, we can safely restrict to posteriors in $\mathcal{K}$. Since there are $|\mathcal{K}| = $ poly $\left(d^{\log(\bar{n}\varrho/\epsilon)/\epsilon^2} \right)$ posteriors, the following linear program (LP 10) has $O(|\mathcal{K}|)$ variables and constraints, and finds an $\alpha(1 - \delta)$-approximation of the optimal signaling scheme:

$$\max_{\gamma \in \Delta_\mathcal{K}} \sum_{\mathbf{p} \in \mathcal{K}} \gamma_{\mathbf{p}} \sum_{\theta \in \Theta} p_\theta \, f_\theta(g(\mathbf{p}, \mathcal{M}_\epsilon(\mathbf{p}))) \tag{10a}$$

$$\text{s.t.} \sum_{p \in \mathcal{K}} \gamma_p \, p_\theta = \mu_\theta \qquad \forall \theta \in \Theta \tag{10b}$$

Given the distribution on the k-uniform posteriors γ, we can construct a direct signaling scheme ϕ by setting:

$$\phi_\theta(\mathbf{a}) = \sum_{\mathbf{p} \in \mathcal{K}: \mathbf{a} = g(\mathbf{p}, \mathcal{M}_\epsilon(\mathbf{p}))} \gamma_{\mathbf{p}} p_\theta, \text{ for each } \theta \in \Theta \text{ and } \mathbf{a} \in \mathcal{A}.$$

This shows that such signaling scheme ϕ is $\alpha(1 - \delta)$-approximate and ϵ-persuasive, which are precisely our desiderata, thus concluding the proof. □

Declarations

Appendix A Omitted Proofs

A.1 Omitted Proof for Lemma 3

Lemma 3 *Given a 3SAT formula φ, if φ is unsatisfiable, then for each (possibly randomized) Merlin$_2$'s strategy η_2 there exists a set S_i such that each Merlin$_1$'s assignment to variables in S_i is in conflict with Merlin$_2$'s assignment with a probability of at least $\rho/2v$.*

Proof Let $\omega(\mathcal{F}_\varphi, \eta_2|S_i)$ be the probability with which Arthur accepts Merlin's answers when Merlin$_1$ receives S_i, and Merlin$_2$ follows strategy η_2. Formally:

$$\omega(\mathcal{F}_\varphi, \eta_2|S_i) := \max_{\eta_1} \mathbb{E}_{T_i}[\mathcal{V}(S_i, T_i, \eta_1, \eta_2)].$$

By definition of the value of a free game, we have:

$$\omega(\mathcal{F}_\varphi) = \frac{1}{m} \max_{\eta_2} \sum_{S_i} \omega(\mathcal{F}_\varphi, \eta_2|S_i) \geq \max_{\eta_2} \min_{S_i} \omega(\mathcal{F}_\varphi, \eta_2|S_i).$$

Then, by Lemma 2, this results in:

$$\max_{\eta_2} \min_{S_i} \omega(\mathcal{F}_\varphi, \eta_2|S_i) \leq 1 - \frac{\rho}{2v},$$

which proves the statement of the lemma. $\qquad\square$

A.2 Omitted Proofs for Theorem 7

Lemma 5 *If $||\mathbf{v}_1 - \tilde{\mathbf{v}}||_1 > 16\epsilon$, there exists a row i of type $\mathtt{t}_2$ such that $w_i < -\epsilon$.*

Proof Lemma 4 implies that, if $||\mathbf{v}_1 - \tilde{\mathbf{v}}||_1 > 16\epsilon$, then there exists a subset $\mathcal{T} \subseteq \{T_j\}_{j\in[m]}$ such that

$$\sum_{T_j\in\mathcal{T}} \sum_{\xi_2} x_{T_j,\xi_2}$$
$$> (1/2 + 4\epsilon) \sum_{j,\xi_2} x_{T_j,\xi_2}$$
$$> 1/4 + \epsilon.$$

It follows that

$$\sum_{T_j\notin\mathcal{T}} \sum_{\xi_2} x_{T_j,\xi_2} < 1/2 + \epsilon - 1/4 - \epsilon = 1/4,$$

which implies that row $(\mathtt{t}_2, \mathcal{T})$ is such that $w_{\mathtt{t}_2,\mathcal{T}} < -1/4 - \epsilon + 1/4 < -\epsilon$. $\qquad\square$

Lemma 6 *If $||\mathbf{v}_2 - \tilde{\mathbf{v}}||_1 > 16\epsilon$, there exists a row i of type t_3 such that $w_i < -\epsilon$.*

Proof Lemma 4 implies that, if $||\mathbf{v}_2 - \tilde{\mathbf{v}}||_1 > 16\epsilon$, then there exists a set $\mathcal{I} \subseteq [8m]$ such that

$$\sum_{i \in \mathcal{I}} \sum_{\ell} x_{i,\ell}$$
$$> (1/2 + 4\epsilon) \sum_{i,\ell} x_{i,\ell}$$
$$> 1/4 + \epsilon.$$

Then,

$$\sum_{i \notin \mathcal{I}} \sum_{\ell} x_{i,\ell} < 1/2 + \epsilon/2 - 1/4 - \epsilon = 1/4 - \epsilon/2.$$

It follows that there exists a row $(\mathsf{t}_3, \mathcal{I})$ such that $w_{\mathsf{t}_3,\mathcal{I}} < -1/4 - \epsilon + 1/4 - \epsilon/2 < -\epsilon$. □

Lemma 7 *For each S_i, $i \in [m]$, there exists at most one row $(\mathsf{t}_4, S_i, \xi_1)$ such that $w_{(\mathsf{t}_4, S_i, \xi_1)} \geq -\epsilon$.*

Proof Let

$$\mathfrak{f}(\mathbf{x}, \xi_1) := \sum_{j:\ell_i = e(\xi_1)_j} x_{j,\ell}.$$

Assume, by contradiction, that for a given S_i there exist two assignments ξ_1' and ξ_1'' such that $w_{(\mathsf{t}_4, S_i, \xi_1)} \geq -\epsilon$ for each $\xi_1 \in \{\xi_1', \xi_1''\}$. Then, $\mathfrak{f}(\mathbf{x}, \xi_1) \geq 1/2 - \epsilon$, for each $\xi_1 \in \{\xi_1', \xi_1''\}$. Otherwise, we would get $w_{(\mathsf{t}_4, S_i, \xi_1)} < 1/2(1/2 - \epsilon) - 1/2(1/2 + \epsilon) = -\epsilon$ for at least one $\xi_1 \in \{\xi_1', \xi_1''\}$. Let $\mathbf{x}'$ be the vector such that

$$x_{i,\ell}' := \frac{x_{i,\ell}}{\sum_{i,\ell} x_{i,\ell}}.$$

Then, for $\xi_1 \in \{\xi_1', \xi_1''\}$,

$$\mathfrak{f}(\mathbf{x}', \xi_1) \geq \frac{1/2 - \epsilon}{1/2 + \epsilon} \geq 1 - 4\epsilon.$$

By Lemmas 4 and 6, we have that $||\mathbf{v}_2 - \tilde{\mathbf{v}}||_1 \leq 16\epsilon$. Therefore, we can obtain a uniform vector $\tilde{\mathbf{x}}$ by moving at most 16ϵ probability from $\mathbf{x}'$. This results in a decrease of $\mathfrak{f}$ of at most 16ϵ, that is $\mathfrak{f}(\tilde{\mathbf{x}}, \xi_1) \geq 1 - 20\epsilon$ for each $\xi_1 \in \{\xi_1', \xi_1''\}$.

By construction $\text{dist}(e) = 1/5$, which implies $\text{dist}(e(\xi_1'), e(\xi_1'')) \geq 1/5$. Then, there exists a set of indices $\mathcal{I}$, with $|\mathcal{I}| \geq 8m/5$, such that $e(\xi_1')_j \neq e(\xi_1'')_j$ for each $j \in \mathcal{I}$.

Therefore,

$$\mathfrak{f}(\tilde{\mathbf{x}}, \xi'_1) + \mathfrak{f}(\tilde{\mathbf{x}}, \xi''_1) \leq \sum_{j \in \mathcal{I}} 1/8m + \sum_{j \notin \mathcal{I}} 2/8m \leq 2 - 1/5.$$

This leads to a contradiction with $\mathfrak{f}(\tilde{\mathbf{x}}, \xi'_1) + \mathfrak{f}(\tilde{\mathbf{x}}, \xi''_1) \geq 2 - 40\epsilon$. $\qquad\square$

A.3 Omitted Proofs for Theorem 2

Lemma 9 *Given* $\mathbf{p}^* \in \Delta_\Theta$, *for each* $\theta \in \Theta$, *and for each* $i \in [k]$ *such that*

$$\left| \frac{i}{k} - p_\theta^* \right| \leq \frac{\epsilon}{4},$$

it holds:

$$\sum_{\mathbf{p} \in \mathcal{P}: p_\theta = i/k} \gamma_{\mathbf{p}} \geq \left(1 - \frac{\delta}{2} \right) \sum_{\mathbf{p} \in \mathcal{K}: p_\theta = i/k} \gamma_{\mathbf{p}},$$

where γ *is the distribution of* k *i.i.d. samples from* $\mathbf{p}^*$.

Proof Fix $\bar{\theta} \in \Theta$ and $i \in [k]$ with $|i/k - p_{\bar{\theta}}^*| \leq \epsilon/4$. Then, for each $r \in \mathcal{R}$ and $a \in \mathcal{A}^r$, let $\tilde{t}_a^r := \sum_\theta \tilde{\gamma}_\theta^{(\bar{\theta}, i)} u_\theta^r(a)$ and $t_a^r := \sum_\theta p_\theta^* u_\theta^r(a)$.
First, we show that $|\mathbb{E}[\tilde{t}_a^r] - t_a^r| \leq \epsilon/4$.
Equivalently,

$$\left| \sum_\theta u_\theta^r(a) \left(\mathbb{E}[\tilde{\gamma}_\theta^{(\bar{\theta}, i)}] - p_\theta^* \right) \right| \leq \frac{\epsilon}{4}.$$

Assume $i/k \geq p_{\bar{\theta}}^*$.
Then,

$$\sum_\theta |\mathbb{E}[\tilde{\gamma}_\theta^{(\bar{\theta}, i)}] - p_\theta^*| = \frac{i}{k} - p_{\bar{\theta}}^* + \sum_{\theta \neq \bar{\theta}} \left(p_\theta^* - \frac{p_\theta^*}{\sum_{\theta' \neq \bar{\theta}} p_{\theta'}^*} \cdot \left(1 - \frac{i}{k} \right) \right)$$

$$\leq \frac{\epsilon}{4} + 1 - p_{\bar{\theta}}^* - 1 + \frac{i}{k}$$

$$\leq \frac{\epsilon}{2}. \qquad\qquad\qquad\qquad\qquad (A1)$$

Analogously, if $i/k \leq p_{\bar{\theta}}^*$, we get that

$$\sum_\theta \left| \mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta}, i)} \right] - p_\theta^* \right| \leq \frac{\epsilon}{2}.$$

Furthermore, let

$$M_1 := \left\{ \theta \in \Theta : \mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \geq 0 \right\},$$

and $M_2 := \Theta \setminus M_1$.

Then,

$$\sum_{\theta \in M_1} \left(\mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \right) = - \sum_{\theta \in M_2} \left(\mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \right) \leq \frac{\epsilon}{4}, \qquad \text{(A2)}$$

where the equality comes from $\sum_\theta \mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] = \sum_\theta p_\theta^* = 1$ and the inequality follows from Eq. A1. Then,

$$\sum_\theta u_\theta^r(a) \left(\mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \right)$$
$$= \sum_{\theta \in M_1} u_\theta^r(a) \left(\mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \right) + \sum_{\theta \in M_2} u_\theta^r(a) \left(\mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \right)$$
$$\leq \frac{\epsilon}{4},$$

where we use both

$$\sum_{\theta \in M_2} u_\theta^r(a) \left(\mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \right) \leq 0$$

and

$$\sum_{\theta \in M_1} u_\theta^r(a) \left(\mathbb{E}\left[\tilde{\gamma}^{(\bar{\theta},i)} \right] - p_\theta^* \right) \leq \frac{\epsilon}{4}$$

by Equation (A2).

Analogously, it is possible to show that

$$\sum_\theta u_\theta^r(a) \left(\mathbb{E}\left[\tilde{\gamma}_\theta^{(\bar{\theta},i)} \right] - p_\theta^* \right) \geq -\frac{\epsilon}{4}.$$

Then,

$$\Pr(|t_a^r - \tilde{t}_a^r| \geq \epsilon/2) \leq \Pr(|\tilde{t}_a^r - \mathbb{E}[\tilde{t}_a^r]| \geq \epsilon/4).$$

Moreover, by the Hoeffding's inequality, we have that, for each $r \in \mathcal{R}$ and $a \in \mathcal{A}^r$,

$$\Pr(|\tilde{t}_a^r - \mathbb{E}[\tilde{t}_a^r]| \geq \epsilon/4) \leq 2e^{-2k(\frac{\epsilon}{4})^2} = 2e^{-4\epsilon^2 \log(4\bar{n}\varrho/\delta)/\epsilon^2} = 2\left(\frac{\delta}{4\bar{n}\varrho} \right)^4 \leq \frac{\delta}{2\bar{n}\varrho}.$$

The union bound yields the following:

$$
\Pr\left(\bigcap_{r\in\mathcal{R},a\in\mathcal{A}^r}|\tilde{t}_a^r-t_a^r|\leq\frac{\epsilon}{2}\right)\geq 1-\sum_{r,a}\Pr\left(|\tilde{t}_a^r-t_a^r|\geq\frac{\epsilon}{2}\right)
$$

$$
\geq\ 1-\sum_{r,a}\Pr\left(|\tilde{t}_a^r-\mathbb{E}[\tilde{t}_a^r]|\geq\frac{\epsilon}{4}\right)
$$

$$
=1-\frac{\delta}{2}.
$$

By the definition of $\mathcal{P}$, this implies that $\Pr(\tilde{\gamma}^{(\bar{\theta},i)}\in\mathcal{P})\geq 1-\delta/2$.
Finally,

$$
\sum_{\mathbf{p}\in\mathcal{P}:p_{\bar{\theta}}=i/k}\gamma_p\Pr\left(\tilde{\gamma}_{\bar{\theta}}=\frac{i}{k}\right)\cdot\Pr\left(\tilde{\gamma}\in\mathcal{P}\mid\tilde{\gamma}_{\bar{\theta}}=\frac{i}{k}\right)
$$

$$
=\Pr\left(\tilde{\gamma}_{\bar{\theta}}=\frac{i}{k}\right)\cdot\Pr\left(\tilde{\gamma}^{(\bar{\theta},i)}\in\mathcal{P}\right)
$$

$$
\geq\left(1-\frac{\delta}{2}\right)\Pr\left(\tilde{\gamma}_{\bar{\theta}}=\frac{i}{k}\right)
$$

$$
=\left(1-\frac{\delta}{2}\right)\sum_{\mathbf{p}\in\mathcal{K}:p_{\bar{\theta}}=i/k}\gamma_p.
$$

This concludes the proof. □

Lemma 10 *Given* $\mathbf{p}^*\in\Delta_\Theta$, *for each* $\theta\in\Theta$, *it holds:*

$$
\sum_{i:|i/k-\mathbf{p}_\theta^*|\geq\epsilon/4}\ \sum_{\mathbf{p}\in\mathcal{K}:p_\theta=i/k}\gamma_p\leq\frac{\delta}{2}\,p_\theta^*,
$$

where γ *is the distribution of* k *i.i.d. samples from* $\mathbf{p}^*$.

Proof The random variable $\tilde{\gamma}_\theta$ is drawn from a binomial distribution.
We consider three possible cases. If $p_\theta^*\geq 1/8$, then, by Hoeffding's inequality, we have

$$
\Pr\left(|\tilde{\gamma}_\theta-p_\theta^*|\geq\frac{\epsilon}{4}\right)\leq 2e^{-2k(\epsilon/4)^2}
$$

$$
=2e^{-4\log(4\bar{n}\varrho/\delta)}
$$

$$
\leq\frac{\delta}{2}\,p_\theta^*.
$$

If $p_\theta^* \leq 1/8$, then, by Chernoff's bound, it holds

$$\Pr\left(\tilde{\gamma}_\theta - p_\theta^* \geq \frac{\epsilon}{4}\right) \leq e^{-k\left(\frac{\epsilon}{4}\right)^2 \frac{1}{1-2p_\theta^*}\log\left(\frac{1-p_\theta^*}{p_\theta^*}\right)}$$

$$\leq e^{-2\log\left(\frac{4\tilde{n}\varrho}{\delta}\right)\log\left(\frac{7}{8p_\theta^*}\right)}$$

$$\leq \left(\frac{8}{7}p_\theta^*\right)^{2\log\left(\frac{4}{\delta}\right)} \tag{A3}$$

$$\leq \frac{8}{7}e\,p_\theta^*\,e^{-2\log\left(\frac{4}{\delta}\right)}$$

$$\leq \frac{\delta}{4}p_\theta^*, \tag{A4}$$

where, to get from (A3) to (A4), we use the fact that $8/7\,e\,p_\theta^* \leq 1$ and $2\log(4/\delta) \geq 1$. Moreover,

$$\Pr\left(\tilde{\gamma}_\theta - p_\theta^* \leq -\frac{\epsilon}{4}\right) \leq e^{-k(\epsilon/4)^2\frac{1}{2(1-p_\theta^*)p_\theta^*}}$$

$$\leq e^{-\frac{\log(4\tilde{n}\varrho/\delta)}{p_\theta^*}} = \left(e^{\frac{1}{p_\theta^*}}\right)^{\log\left(\frac{\delta}{4}\right)}$$

$$\leq \left(\frac{1}{p_\theta^*}e\right)^{\log\left(\frac{\delta}{4}\right)} \tag{A5}$$

$$\leq \left(\frac{1}{p_\theta^*}\right)^{-1}e^{\log\left(\frac{\delta}{4}\right)} \tag{A6}$$

$$= \frac{\delta}{4}p_\theta^*.$$

where in (A5) we use the fact that $e^x \geq ex$, and in (A6) that $\log(\delta/4) < -1$.

Then,

$$\sum_{i:|i/k-\mathbf{p}_\theta^*|\geq\epsilon/4}\ \sum_{\mathbf{p}\in\mathcal{K}:p_\theta=i/k} \gamma_p = \Pr\left(|\tilde{\gamma}_\theta - p_\theta^*| \geq \frac{\epsilon}{4}\right) \leq \frac{\delta}{2}p_\theta^*,$$

which concludes the proof of the lemma. $\qquad\square$

Lemma 11 *Given a* $\mathbf{p}^* \in \Delta_\Theta$, *for each* $\theta \in \Theta$, *it holds:*

$$\sum_{\mathbf{p}\in\mathcal{P}}\gamma_p\,p_\theta \geq (1-\delta)\,p_\theta^*,$$

where γ *is the distribution of* k *i.i.d. samples from* $\mathbf{p}^*$.

Proof We show the following:

$$\sum_{\mathbf{p}\in\mathcal{P}} \gamma_{\mathbf{p}} p_\theta \geq \sum_{i:|i/k-p_\theta^*|\leq\epsilon/4} \frac{i}{k} \sum_{\mathbf{p}\in\mathcal{P}:p_\theta=i/k} \gamma_p \tag{A7}$$

$$\geq \sum_{i:|i/k-p_\theta^*|\leq\epsilon/4} \frac{i}{k} \sum_{\mathbf{p}\in\mathcal{K}:p_\theta=i/k} \left(1-\frac{\delta}{2}\right)\gamma_{\mathbf{p}} \tag{A8}$$

$$= \left(1-\frac{\delta}{2}\right) \sum_{i:|i/k-p_\theta^*|\leq\epsilon/4} \frac{i}{k} \sum_{\mathbf{p}\in\mathcal{K}:p_\theta=i/k} \gamma_{\mathbf{p}}$$

$$\geq \left(1-\frac{\delta}{2}\right)\left(p_\theta^* - \sum_{i:|i/k-p_\theta^*|\geq\epsilon/4} \frac{i}{k} \sum_{\mathbf{p}\in\mathcal{K}:\mathbf{p}_\theta=i/k} \gamma_{\mathbf{p}}\right)$$

$$\geq \left(1-\frac{\delta}{2}\right)\left(p_\theta^* - \sum_{i:|i/k-p_\theta^*|\geq\epsilon/4} \sum_{\mathbf{p}\in\mathcal{K}:\mathbf{p}_\theta=i/k} \gamma_{\mathbf{p}}\right) \tag{A9}$$

$$\geq \left(1-\frac{\delta}{2}\right)^2 p_\theta^*$$

$$\geq (1-\delta)\, p_\theta^*. \tag{A10}$$

Equation (A7) holds since we are restricting the set of posteriors; Eq. (A8) holds by Lemma 9; Eq. (A9) holds since $i/k \leq 1$; and Eq. (A10) holds by Lemma 10. This concludes the proof of the lemma. □

Lemma 12 *Given* $\mathbf{p}^* \in \Delta_\Theta$, *for each* $\mathbf{p} \in \mathcal{P}$, *it holds:* $\mathcal{M}(\mathbf{p}^*) \subseteq \mathcal{M}_\epsilon(\mathbf{p})$.

Proof Let $Z_1 = \mathcal{M}_\epsilon(\mathbf{p})$ and $Z_2 = \mathcal{M}(\mathbf{p}^*)$, and $a \in Z_2^r$. Then, for all $a' \in \mathcal{A}^r$,

$$\sum_\theta p_\theta\, u_\theta^r(a) \geq \sum_\theta p_\theta^* u_\theta^r(a) - \frac{\epsilon}{2} \geq \sum_\theta p_\theta^* u_\theta^r(a') - \frac{\epsilon}{2} \geq \sum_\theta p_\theta\, u_\theta^r(a') - \epsilon.$$

Thus, $a \in Z_1^r$, which proves the lemma. □

Lemma 13 *Given a* $\mathbf{p}^* \in \Delta_\Theta$, *it holds:*

$$\sum_{\mathbf{p}\in\mathcal{K}} \gamma_{\mathbf{p}} \sum_\theta p_\theta\, f_\theta(g(\mathbf{p}, \mathcal{M}_\epsilon(\mathbf{p}))) \geq \alpha(1-\delta) f_\theta(g^*(\mathbf{p}^*, \mathcal{M}(\mathbf{p}^*))),$$

where γ *is the distribution of* k *i.i.d. samples from* $\mathbf{p}^*$.

Proof We prove the following:

$$\sum_{\mathbf{p} \in \mathcal{K}} \gamma_{\mathbf{p}} \sum_{\theta} p_\theta \, f_\theta(g(\mathbf{p}, \mathcal{M}_\epsilon(\mathbf{p}))) \tag{A11}$$

$$\geq \alpha \sum_{\mathbf{p} \in \mathcal{K}} \gamma_{\mathbf{p}} \sum_{\theta} p_\theta \, f_\theta(g^*(\mathbf{p}, \mathcal{M}_\epsilon(\mathbf{p}))) \tag{A12}$$

$$\geq \alpha \sum_{\mathbf{p} \in \mathcal{P}} \gamma_{\mathbf{p}} \sum_{\theta} p_\theta \, f_\theta(g^*(\mathbf{p}, \mathcal{M}_\epsilon(\mathbf{p}))) \tag{A13}$$

$$\geq \alpha \sum_{\mathbf{p} \in \mathcal{P}} \gamma_{\mathbf{p}} \sum_{\theta} p_\theta \, f_\theta(g^*(\mathbf{p}^*, \mathcal{M}_\epsilon(\mathbf{p}))) \tag{A14}$$

$$\geq \alpha \sum_{\mathbf{p} \in \mathcal{P}} \gamma_{\mathbf{p}} \sum_{\theta} p_\theta \, f_\theta(g^*(\mathbf{p}^*, \mathcal{M}(\mathbf{p}^*))) \tag{A15}$$

$$\geq \alpha \, (1 - \delta) \sum_{\theta} p_\theta^* \, f_\theta(g^*(\mathbf{p}^*, \mathcal{M}(\mathbf{p}^*))). \tag{A16}$$

Equation (A11) is the relaxed sender's expected utility; Eq. (A12) holds by Definition 8; Eq. (A13) holds by restricting the set of posteriors; Eq. (A14) holds by the optimality of g^*; Eq. (A15) holds by Lemma 12; and Eq. (A16) holds by Lemma 11. This concludes the proof. ☐

References

1. Badanidiyuru, A., Bhawalkar, K., Xu, H.: Targeting and signaling in ad auctions. In: SODA, pp. 2545–2563 (2018)
2. Emek, Y., Feldman, M., Gamzu, I., PaesLeme, R., Tennenholtz, M.: Signaling schemes for revenue maximization. ACM Trans. Econ. Comput. **2**(2), 1–19 (2014)
3. Miltersen, P.B., Sheffet, O.: Send mixed signals: earn more, work less. In: ACM EC, pp. 234–247 (2012)
4. Alonso, R., Câmara, O.: Persuading voters. Am. Econ. Rev. **106**(11), 3590–3605 (2016)
5. Cheng, Y., Cheung, H.Y., Dughmi, S., Emamjomeh-Zadeh, E., Han, L., Teng, S.: Mixture selection, mechanism design, and signaling. In: FOCS, pp. 1426–1445 (2015)
6. Bhaskar, U., Cheng, Y., Ko, Y.K., Swamy, C.: Hardness results for signaling in Bayesian zero-sum and network routing games. In: ACM EC, pp. 479–496 (2016)
7. Vasserman, S., Feldman, M., Hassidim, A.: Implementing the wisdom of waze. In: IJCAI (2015)
8. Mansour, Y., Slivkins, A., Syrgkanis, V., Wu, Z.S.: Bayesian exploration: Incentivizing exploration in Bayesian games. In: ACM EC, pp. 661–661 (2016)
9. Rabinovich, Z., Jiang, A.X., Jain, M., Xu, H.: Information disclosure as a means to security. In: AAMAS, pp. 645–653 (2015)
10. Xu, H., Freeman, R., Conitzer, V., Dughmi, S., Tambe, M.: Signaling in Bayesian Stackelberg games. In: AAMAS, pp. 150–158 (2016)
11. Xu, H., Rabinovich, Z., Dughmi, S., Tambe, M.: Exploring information asymmetry in two-stage security games. In: AAAI (2015)
12. Babichenko, Y., Barman, S.: Algorithmic aspects of private Bayesian persuasion. In: Innovations in Theoretical Computer Science Conference (2017)
13. Candogan, O.: Persuasion in networks: Public signals and k-cores. Available at SSRN (2019)
14. Kamenica, E., Gentzkow, M.: Bayesian persuasion. Am. Econ. Rev. **101**(6), 2590–2615 (2011)
15. Dughmi, S.: Algorithmic information structure design: a survey. ACM SIGEC Exch **15**(2), 2–24 (2017)
16. Rayo, L., Segal, I.: Optimal information disclosure. J Polit Econ **118**(5), 949–987 (2010)

17. Castiglioni, M., Gatti, N.: Persuading voters in district-based elections. In: The AAAI Conference on Artificial Intelligence, (AAAI 2021) (2021)
18. Arieli, I., Babichenko, Y.: Private Bayesian persuasion. J. Econ. Theory **182**, 185–217 (2019)
19. Rubinstein, A.: Honest Signaling in Zero-Sum Games Is Hard, and Lying Is Even Harder. In: 44th International Colloquium on Automata, Languages, and Programming (ICALP 2017), pp. 77–17713 (2017)
20. Babichenko, Y., Barman, S.: Algorithmic aspects of private Bayesian persuasion. In: 8th Innovations in Theoretical Computer Science Conference (ITCS 2017) (2017)
21. Dughmi, S., Xu, H.: Algorithmic persuasion with no externalities. In: ACM EC, pp. 351–368 (2017)
22. Xu, H.: On the tractability of public persuasion with no externalities. In: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 2708–2727 (2020). SIAM
23. Dughmi, S., Xu, H.: Algorithmic Bayesian persuasion. In: ACM STOC, pp. 412–425 (2016)
24. Castiglioni, M., Celli, A., Gatti, N.: Persuading voters: It's easy to whisper, it's hard to speak loud. In: Proceedings of the AAAI Conference on Artificial Intelligence vol. 34, pp. 1870–1877 (2020)
25. Amaldi, E., Kann, V.: The complexity and approximability of finding maximum feasible subsystems of linear relations. Theor. Comput. Sci. **147**(1&2), 181–210 (1995)
26. Daskalakis, C., De, A., Diakonikolas, I., Moitra, A., Servedio, R.: A polynomial-time approximation scheme for fault-tolerant distributed storage. In: SODA, pp. 628–644 (2014)
27. Aaronson, S., Impagliazzo, R., Moshkovitz, D.: Am with multiple merlins. In: CCC, pp. 44–55 (2014)
28. Deligkas, A., Fearnley, J., Savani, R.: Inapproximability results for approximate nash equilibria. In: International Conference on Web and Internet Economics, pp. 29–43 (2016). Springer
29. Bergemann, D., Morris, S.: Information design: a unified perspective. J. Econ. Lit. **57**(1), 44–95 (2019)
30. Kamenica, E.: Bayesian persuasion and information design. Annu. Rev. Econ. **11**, 249 (2018)
31. Bergemann, D., Morris, S.: Bayes correlated equilibrium and the comparison of information structures in games. Theor. Econ. **11**(2), 487–522 (2016)
32. Bergemann, D., Morris, S.: Information design, Bayesian persuasion, and Bayes correlated equilibrium. Am. Econ. Rev. **106**(5), 586–91 (2016)
33. Brandt, F., Conitzer, V., Endriss, U., Lang, J., Procaccia, A.D.: Handbook of computational social choice. Cambridge University Press (2016)
34. Babichenko, Y., Papadimitriou, C., Rubinstein, A.: Can almost everybody be almost happy? In: Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, pp. 1–9 (2016)
35. Gilbert, E.N.: A comparison of signalling alphabets. Bell Syst. Tech. J. **31**(3), 504–522 (1952)
36. Dinur, I.: The PCP theorem by gap amplification. J. ACM **54**(3), 12 (2007)
37. Even, S., Selman, A.L., Yacobi, Y.: The complexity of promise problems with applications to public-key cryptography. Inf. Control **61**(2), 159–173 (1984)
38. Goldreich, O.: On promise problems: A survey. In: Theoretical Computer Science, pp. 254–290 (2006)