Selling Information While Being an Interested Party

Francesco Bacchiocchi*
Politecnico di Milano
francesco.bacchiocchi@polimi.it

Matteo Castiglioni*
Politecnico di Milano
matteo.castiglioni@polimi.it

Nicola Gatti*
Politecnico di Milano
nicola.gatti@polimi.it

Alberto Marchesi*
Politecnico di Milano
alberto.marchesi@polimi.it

Giulia Romano*
Politecnico di Milano
giulia.romano@polimi.it

Abstract

We study the algorithmic problem faced by an information holder (*seller*) who *wants to optimally sell information* to a budget-constrained decision maker (*buyer*). The utilities of both agents depend on a random *state of nature* that is revealed to the seller, but unknown to the buyer. Differently from previous works, we consider the case in which the seller is an *interested party*, as the decision taken by the buyer also influences seller's utility. The seller's goal is to (partially) sell their information about the state of nature to the buyer, so as to concurrently maximize revenue and induce the buyer to take a desirable decision. We study settings in which buyer's budget and utilities are determined by a random buyer's *type* unknown to the seller. First, we propose a polynomial-time algorithm for computing an *optimal* seller's protocol, which proposes a *menu* of information-revelation policies to the buyer, who acquires one of them by paying its corresponding *price*. Then, we switch the attention to the case in which the seller can only employ a single information-revelation policy, rather than proposing a menu. In such a setting, we give hardness and approximation results for the seller's algorithmic problem.

1 Introduction

Nowadays, a large amount of information is collected on the Web and other online platforms. Such information ranges from consumer preferences, *e.g.*, in e-commerce and streaming websites, to credit reports and location histories. As a result, recent years have seen the emergence and rapid growth of markets where specialized companies sell information valuable to other businesses, such as advertisers, retailers, and loan providers. Very recently, information markets have also attracted the attention of the algorithmic game theory community. However, while works on classical settings such as auctions [13], signaling [14], and contract design [16] are now proliferating, only a few papers study *information selling*, with [4] and [12] as notable examples (see Section A).

We study the algorithmic problem faced by an information holder (*seller*) who *wants to optimally sell information* to a budget-constrained decision maker (*buyer*) that must take an action. Differently from previous works (see, *e.g.*, [12]), we consider a setting where the seller is an *interested party*, since the buyer's action affects not only their own utility, but also that of the seller. This arises in many real-world settings, where the use of acquired information may generate externalities that impact the

*Authors are listed in alphabetical order.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

seller. Both agents' utilities depend on a *state of nature* drawn from a commonly-known distribution, observed by the seller but not by the buyer. The seller's goal is then to (partially) sell this information so as to maximize revenue while inducing a desirable action by the buyer.

We study settings in which the buyer's budget and utilities are determined by the buyer's random *type*, which is unknown to the seller. In such settings, to optimally sell information, the seller has to commit upfront to a *protocol* working as follows. First, the seller proposes to the buyer a *menu* of information-revelation policies, and the buyer selects an expected-utility-maximizing one according to their private type, paying its corresponding *price*. By building on the *Bayesian persuasion* framework of [22], an information-revelation policy is implemented as a *signaling scheme*, i.e., a randomized mapping from states of nature to signals sent to the buyer. Then, the realized state of nature is disclosed to the seller, who reveals information about it to the buyer according to the acquired signaling scheme. Finally, the buyer selects a best-response action according to the just acquired information, and the seller pays back the buyer with a *payment* that depends on both the chosen action and the signal previously sent by the seller. Our protocol extends the one of Chen et al. [12] by adding a final payment from the seller to the buyer. As we show later, this is crucial in order to design optimal protocols in our setting where the seller is an interested party, since the latter is *not* only concerned with revenue, but also with the buyer's action. Moreover, the addition of payments from the seller to the buyer is also reasonable in many real-world scenarios. For instance, think of a case in which the information holder asks the buyer to deposit additional money, and this is given back to them only if the performed action respects some given rules on which the two parties agreed upfront (see Appendix B for additional real-world motivating examples).

1.1 Original Contributions

In Section 3, we prove our first main result, for the case of general protocols *with menus*. This is informally stated as:

Theorem 1.1 (Informal)**.** *A seller-optimal protocol with menus can be computed in polynomial time.*

To prove the result, we first formulate the goal as solving a quadratic problem. Then, we show that one can restrict attention to *direct* and *persuasive* signaling schemes, which send signals corresponding to action recommendations and properly incentivize the buyer to follow them. This allows us to focus on protocols that require the buyer to pay their entire budget upfront and refund them only if they take the recommended action. These results allow us to derive a suitable linear relaxation of the quadratic problem. A similar technique has been used in principal-agent problems [19, 7], where an optimal solution to the linear relaxation can be efficiently cast into an approximate solution to the quadratic problem. Surprisingly, in our information-selling setting, an optimal solution to the linear relaxation allows us to efficiently recover an optimal protocol.

In the second part of the paper, we switch the attention to the case of protocols *without menus*, where the seller can only commit to a single signaling scheme. This is the case in many real-world applications, where it is unreasonable that a buyer is asked to choose among a range of options. Computing an optimal protocol without menus begets considerable additional challenges, since the seller has no way of extracting information about the buyer's private type.

In Section 4, we draw a connection between computing an optimal protocol without menus and principal-agent problems with *observable actions*. In such problems, a principal commits to an action-dependent payment scheme so as to incentivize an agent to take a costly observable action. We prove that these problems are a special case of ours, and that computing a utility-maximizing payment-scheme for the principal is APX-hard. These results show that our information-selling problem is APX-hard even when the number of states of nature is fixed and the buyer has *limited liability*, *i.e.*, they have *no* budget at their disposal.

In Section 5, we circumvent the APX-hardness result when the seller employs protocols without menus and the buyer has *limited liability*. This case is of interest on its own, as similar models have been addressed by [15] in signaling. Our results can be informally stated as:

Theorem 1.2 (Informal)**.** *In information-selling problems in which the buyer has limited liability, the problem of computing a seller-optimal protocol without menus admits: (i) an additive PTAS when the number of buyer's actions is fixed; (ii) a polynomial-time bi-criteria approximation algorithm when the number of states of nature is fixed.*

Table 1: Each cell specifies the computational complexity of finding an optimal protocol and the approximation guarantee achievable in polynomial time. "*bi-criteria*" stands for $\rho\text{OPT} - 2^{-\Omega(1/\rho)} - \epsilon$, where OPT is the optimal seller's expected utility. The approximation guarantees that are shaded in gray can only be obtained in quasi-polynomial time.

	general	fixed #actions	fixed #states	fixed #types
Prot. *w. menus*	P	P	P	P
Prot. *w/o menus* *w. limited liability*	APX-hard	—	APX-hard	P
	bi-criteria	additive PTAS	*bi-criteria*	
Prot. *w/o menus* *w/o limited liability*	APX-hard	no PTAS	APX-hard	P
	bi-criteria	*bi-criteria*	*bi-criteria*	

Our *bi-criteria* approximation algorithms take any $\rho \in (0, 1/2]$ and $\epsilon > 0$ as inputs, and they provide a multiplicative approximation ρ of the optimal seller's expected utility by only suffering a $2^{-\Omega(1/\rho)} + \epsilon$ additive loss. Such an additive loss is exponentially small in $1/\rho$, and, thus, it is negligible even for reasonably large values of ρ. As shown by [9], this approximation is tight for hidden-action principal-agent problems. It remains open whether it is also tight for principal-agent problems with observable actions, a special case of our information-selling problem.

Finally, in Section 6 we study the problem of computing optimal protocols without menus in general settings in which the buyer does *not* have limited liability. We first prove a negative result stronger than the one in Section 4, by showing that, in such a setting, the problem is NP-hard to approximate within a constant factor even if the number of buyer's actions is fixed. Then, we show how to circumvent such a negative result:

Theorem 1.3 (Informal)**.** *In general information-selling problems (i.e., without limited liability), the problem of computing a seller-optimal protocol without menus admits a quasi-polynomial-time bi-criteria approximation algorithm. Moreover, when either the number of buyer's actions or that of states of nature is fixed, such an algorithm is poly-time. (See Table 1)*

2 Preliminaries

The information available to the seller is termed *state of nature* and encoded as an element of a finite set $\Theta \coloneqq \{\theta_i\}_{i=1}^d$ of d possible states, while the set of the m actions available to the buyer is $\mathcal{A} \coloneqq \{a_i\}_{i=1}^m$. The buyer is also characterized by a private *type*, which is unknown to the seller and belongs to a finite set $\mathcal{K} \coloneqq \{k_i\}_{i=1}^n$ of n possible types. Each buyer's type $k \in \mathcal{K}$ is characterized by a utility function $u_\theta^k : \mathcal{A} \to [0, 1]$ associated to each state $\theta \in \Theta$ and a budget $b_k \in \mathbb{R}_+$ representing how much they can afford to pay. In our model, the seller's utility is *not* only determined by how much the buyer pays for acquiring information, but it also depends on the buyer's action. Specifically, for every state $\theta \in \Theta$, the seller gets an additional utility contribution determined by a function $u_\theta^s : \mathcal{A} \to [0, 1]$. We assume that both the seller and the buyer know the probability distribution $\mu \in \Delta_\Theta$ according to which the state of nature is drawn, as well as the probability distribution $\lambda \in \Delta_\mathcal{K}$ determining the buyer's type. The state of nature and the buyer's type are independent.[2] We let μ_θ be the probability assigned to state $\theta \in \Theta$, while λ_k is the probability of type $k \in \mathcal{K}$. As in [12], we assume w.l.o.g. that information revelation happens only once. As in *Bayesian* persuasion [22], the seller reveals information to the buyer by committing to a *signaling scheme* ϕ, which is a randomized mapping from states of nature to signals sent to the buyer. Formally, $\phi : \Theta \to \Delta_\mathcal{S}$, where $\mathcal{S}$ is a finite set of signals. We denote by $\phi_\theta \in \Delta_\mathcal{S}$ the probability distribution employed when the state of nature is $\theta \in \Theta$, with $\phi_\theta(s)$ being the probability of sending $s \in \mathcal{S}$.

2.1 Protocols With Menus

A seller's *protocol* is defined as follows. The seller first proposes a *menu* of signaling schemes to the buyer, with each signaling scheme having a *price*. Then, the buyer chooses a signaling scheme and pays its price upfront, before information is revealed.[3] The seller also commits to action-dependent

[2] In this work, given a finite set X, we let Δ_X be the set of all the probability distributions over the elements of X.

[3] Notice that proposing a menu is equivalent to asking the buyer to report their type, as in [12].

payments, which are made by the seller in favor of the buyer after information is revealed and the latter has taken an action. This is in contrast with what happens in the protocol by [12], where there are no action-dependent money transfers. Intuitively, such action-dependent payments are needed in order to incentivize the agent to play an action that is profitable for the seller, and, thus, they are *not* needed in the setting of [12] where the seller's utility function is only determined by how much the buyer pays for acquiring information.[4] Formally:

Definition 2.1. *A* protocol *is a tuple* $\{(\phi^k, p_k, \pi_k)\}_{k\in\mathcal{K}}$*:*

- $\{\phi^k\}_{k\in\mathcal{K}}$ *is a* menu of signaling schemes *defined as* $\phi^k : \Theta \to \Delta_{\mathcal{S}}$*, one for each receiver's type* $k \in \mathcal{K}$*;*
- $\{p_k\}_{k\in\mathcal{K}}$ *is a* menu of prices*, with* $p_k \in \mathbb{R}_+$ *representing the price of signaling scheme* ϕ^k*;*
- $\{\pi_k\}_{k\in\mathcal{K}}$ *is a* menu of payment functions*, defined as* $\pi_k : \mathcal{S} \times \mathcal{A} \to \mathbb{R}_+$ *with* $\pi_k(s, a)$ *encoding how much the seller pays the buyer for action* $a \in \mathcal{A}$ *after selecting* ϕ^k *and receiving signal* $s \in \mathcal{S}$*.*

The seller-buyer interaction is as follows: (i) the seller commits to a protocol $\{(\phi^k, p_k, \pi_k)\}_{k\in\mathcal{K}}$; (ii) the buyer selects a signaling scheme ϕ^k and pays p_k to the seller (with $k \in \mathcal{K}$ possibly different from their true type); (iii) the seller observes the state of nature $\theta \sim \mu$, draws a signal $s \sim \phi^k_\theta$ according to ϕ^k, and sends it to the buyer; (iv) given the signal s, the buyer infers a posterior distribution $\xi^s \in \Delta_\Theta$ over states of nature, where the probability ξ^s_θ of state $\theta \in \Theta$ is computed with the *Bayes* rule, as $\xi^s_\theta := \frac{\mu_\theta\, \phi^k_\theta(s)}{\sum_{\theta'\in\Theta} \mu_{\theta'} \phi^k_{\theta'}(s)}$; (v) given the posterior ξ^s, the buyer selects an action $a \in \mathcal{A}$; and (vi) the seller pays $\pi_k(s, a)$ to the buyer. As in the model by [12], we assume that the seller is committed to following the protocol, while the buyer is *not*, *i.e.*, the buyer is free to leave the interaction at any point. In step (v), after observing a signal $s \in \mathcal{S}$ and computing the posterior ξ^s, the buyer plays a *best response* by choosing an action $a \in \mathcal{A}$ maximizing their expected utility. Formally:

Definition 2.2. *Let* $\epsilon \geq 0$*. For a signal* $s \in \mathcal{S}$*, the posterior* $\xi^s \in \Delta_\Theta$*, and a payment function* $\pi : \mathcal{S} \times \mathcal{A} \to \mathbb{R}_+$*, the* ϵ-best-response set *of a buyer of type* $k \in \mathcal{K}$ *is the set* $\mathcal{B}^{k,\epsilon}_{\xi^s,\pi}$ *of all the actions* $a \in \mathcal{A}$ *such that, for every* $a' \in \mathcal{A}$*, it holds*

$$\sum_{\theta\in\Theta} \xi^s_\theta\, u^k_\theta(a) + \pi(s, a) \geq \sum_{\theta\in\Theta} \xi^s_\theta\, u^k_\theta(a') + \pi(s, a') - \epsilon.$$

We let $b^{k,\epsilon}_{\xi^s,\pi} \in \mathcal{B}^{k,\epsilon}_{\xi^s,\pi}$ *be an* ϵ-best response *played by the buyer, chosen to maximize the seller's expected utility over this set. Best responses and their set are defined for* $\epsilon = 0$*, and they are denoted by omitting* ϵ *from the notation.*[5][6]

The seller's goal is to implement an optimal (*i.e.*, utility-maximizing) protocol $\{(\phi^k, p_k, \pi_k)\}_{k\in\mathcal{K}}$. We focus w.l.o.g. on seller's protocols that are *incentive compatible* (IC) and *individually rational* (IR). Specifically, a seller's protocol is IC if for every pair of buyer's types $k, k' \in \mathcal{K}$ such that $p_{k'} \leq b_k$, we have:

$$\sum_{s\in\mathcal{S},\theta\in\Theta} \mu_\theta \phi^k_\theta(s) \left[u^k_\theta(b^k_{\xi^s,\pi_k}) + \pi_k(s, b^k_{\xi^s,\pi_k})\right] - p_k \geq \sum_{s\in\mathcal{S}} \max_{a\in\mathcal{A}} \sum_{\theta\in\Theta} \mu_\theta \phi^{k'}_\theta(s) \left[u^k_\theta(a) + \pi_{k'}(s, a)\right] - p_{k'},$$

while it is IR if for every type $k \in \mathcal{K}$ and action $a \in \mathcal{A}$:

$$\sum_{\substack{s\in\mathcal{S}\\ \theta\in\Theta}} \mu_\theta \phi^k_\theta(s) \left[u^k_\theta(b^k_{\xi^s,\pi_k}) + \pi_k(s, b^k_{\xi^s,\pi_k})\right] - p_k \geq \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(a).$$

Intuitively, IC incentivizes the buyer to select the ϕ^k corresponding to their true type $k \in \mathcal{K}$, while IR ensures that the buyer gets at least as much utility by acquiring information rather than leaving the protocol and playing an action without information. Then, the seller's expected utility is:

$$\sum_{k\in\mathcal{K}} \lambda_k \left[\sum_{\substack{s\in\mathcal{S}\\ \theta\in\Theta}} \mu_\theta \phi^k_\theta(s) \left[u^s_\theta(b^k_{\xi^s,\pi_k}) - \pi_k(s, b^k_{\xi^s,\pi_k})\right] + p_k\right].$$

[4] In Appendix B, we show that action-dependent payments allow the seller to extract an expected utility that can be arbitrarily larger than what is achievable without them.

[5] We assume that the buyer always breaks ties in favor of the seller when indifferent among multiple best responses.

[6] We will oftentimes work in the space of the distributions over posteriors. In that case, given a posterior $\xi \in \Delta_\Theta$, we abuse notation and write $\mathcal{B}^k_{\xi,\pi}$, $\mathcal{B}^{k,\epsilon}_{\xi,\pi}$, $b^{k,\epsilon}_{\xi,\pi}$, and $b^k_{\xi,\pi}$.

A crucial component of our results is that we can restrict the attention to protocols that are *direct* and *persuasive*. A protocol is direct if $\mathcal{S} = \mathcal{A}$, and it is persuasive whenever playing the recommended action is always a best response for the buyer.

2.2 Protocols Without Menus

In a *protocol without menus*, the seller does *not* propose a menu to the buyer, but commits to a single signaling scheme and payment function.[7] Thus, we denote a protocol without menus as (ϕ, p, π), where ϕ is a signaling scheme, $p \in \mathbb{R}_+$ its price, and π a payment function. The seller-buyer interaction unfolds as in the general case, but step (ii) only involves the payment of $p \in \mathbb{R}_+$ on buyer's part. Some of our results on protocols without menus address the special case in which the buyer has *limited liability*, which means that the buyer has no budget, and, thus, the seller cannot charge $p > 0$ upfront. Formally, this amounts to asking that $b_k = 0$ for all $k \in \mathcal{K}$. Notice that, while such a special case may seem of scarce appeal for selling-information problems, it is indeed interesting on its own, as it is similar to the model studied by [15]. Indeed, our model can be seen as a generalization of the one in [15] to multiple buyer's types. Moreover, when the buyer has *no* limited liability, our model additionally builds on top of that of [15], by adding the possibility for the seller to make the buyer pay before information revelation.

For protocols without menus, IC constraints are *not* needed, while IR constraints are still required to ensure that the buyer is incentivized to acquire information. Given a protocol, only some buyer's types are actually incentivized to participate in it. Formally, a protocol determines a subset $\mathcal{R}_{\phi,p,\pi} \subseteq \mathcal{K}$ of buyer's types that participate such that, for every $k \in \mathcal{R}_{\phi,p,\pi}$: (i) type k has enough budget to buy information, namely $b_k \geq p$; and (ii) the IR constraint is satisfied for type k.[8] Formally, for every $k \in \mathcal{R}_{\phi,p,\pi}$ and $a \in \mathcal{A}$, it must hold:

$$\sum_{s \in \mathcal{S}, \theta \in \Theta} \mu_\theta \phi_\theta(s) \left[u_\theta^k(b^k_{\xi^s,\pi}) + \pi(s, b^k_{\xi^s,\pi}) \right] - p \geq \sum_{\theta \in \Theta} \mu_\theta u_\theta^k(a).$$

Moreover, the seller's expected utility is given by:

$$\sum_{k \in \mathcal{R}_{\phi,p,\pi}} \lambda_k \left[\sum_{\substack{s \in \mathcal{S} \\ \theta \in \Theta}} \mu_\theta \phi_\theta(s) \left[u_\theta^s(b^k_{\xi^s,\pi}) - \pi(s, b^k_{\xi^s,\pi}) \right] + p \right] + \sum_{k \notin \mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_{\theta \in \Theta} \mu_\theta u_\theta^s(b^k_\mu),$$

where $b^k_\xi \in \arg\max_{a \in \mathcal{A}} \sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a)$ is a best response for a buyer's type $k \in \mathcal{K}$ that only considers the posterior $\xi \in \Delta_\Theta$, with ties broken in favor of the seller. Notice that buyer's types $k \notin \mathcal{R}_{\phi,p,\pi}$ are those who do *not* acquire information, and, thus, they play a best response to the probability distribution μ (instead of a posterior).

Finally, for protocols without menus, it is useful to directly work with *distributions over posteriors* induced by signaling schemes [22]. A signaling scheme $\phi : \Theta \to \Delta_\mathcal{S}$ induces a distribution γ over Δ_Θ, which has support $\text{supp}(\gamma) \coloneqq \{\xi^s \mid s \in \mathcal{S}\}$ and it is *consistent*, *i.e.*, it satisfies $\sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \xi_\theta = \mu_\theta$ for $\theta \in \Theta$, where $\gamma_\xi \in [0, 1]$ is the probability that γ assigns to $\xi \in \text{supp}(\gamma)$. Thus, instead of working with signaling schemes, one can w.l.o.g. work with consistent distributions γ over Δ_Θ. [9]

3 Computing Optimal Protocols With Menus

We now focus on protocols with menus. This can be formulated as a quadratic optimization problem consisting in maximizing the seller's expected utility over all the protocols that are IC and IR (see Section 2). Notice that it is generally non-trivial to establish whether such an optimization problem admits a maximum or not, as in principal-agent problems (see, *e.g.*, [19]). However, our results in this section show that our problem indeed admits a maximum. As a first step, we prove that we can focus w.l.o.g. on protocols that are direct and persuasive.

[7] From the point of view of [12], this amounts to having no type reporting stage.

[8] Whenever the expected utility of a buyer's type is the same by participating in the protocol as *not* doing that, we assume that they take the option maximizing the seller's expected utility. If the seller is also indifferent, the buyer participates.

[9] With a slight abuse of notation, we denote a protocol without menus as (γ, p, π), identifying a signaling scheme with its induced distribution over posteriors γ. Similarly, we assume payment functions to be defined over posteriors rather than signals. Formally, $\pi(\xi, a)$ denotes how much the seller pays the buyer when the posterior is $\xi \in \Delta_\Theta$ and they play $a \in \mathcal{A}$.

Lemma 3.1. *Given any IC and IR protocol, one can always recover an IC and IR protocol that is direct and persuasive, and it provides the seller with the same expected utility.*

Lemma 3.1 holds since, if two signals induce the same best response for a buyer's type $k \in \mathcal{K}$, then one can merge the two signals in a single one, recovering a new protocol that achieves the same seller's expected utility. By doing this for every buyer's type until there are no two signals inducing the same best response, we obtain a direct and persuasive protocol without degrading seller's utility. Notice that, since in direct protocols $\mathcal{S} = \mathcal{A}$, whenever we write $\pi_k(a, a')$ for $a, a' \in \mathcal{A}$, action a is the seller's recommendation (signal), while a' is the action played by the buyer.

As a second step, Lemma 3.1 lets us focus w.l.o.g. on protocols that ask for the entire buyer's budget upfront and do *not* pay back deviations from the seller's recommendation. Formally:

Lemma 3.2. *Given an IC, IR, direct, and persuasive protocol*[10]*, it is always possible to recover an IC, IR, direct, and persuasive protocol $\{(\phi^k, \tilde{p}_k, \tilde{\pi}_k)\}_{k \in \mathcal{K}}$ that provides the same seller's expected utility, and, for every type $k \in \mathcal{K}$, satisfies $\tilde{p}_k = b_k$ and $\tilde{\pi}_k(a, a') = 0$ for all $a \neq a' \in \mathcal{A}$.*

Thus, we compactly denote $\pi_k(a, a)$ as $\pi_k(a)$ for all $a \in \mathcal{A}$. In the following, we introduce an LP that is a linear relaxation of the original quadratic problem. To formulate the LP, we exploit Lemmas 3.1 and 3.2. Moreover, we encode the terms $\sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) \pi_k(a)$ as single variables $l_k(a)$.

$$\max_{\phi, l, y \geq 0} \sum_{k \in \mathcal{K}} \lambda_k \left\{ \sum_{a \in \mathcal{A}} \left[\sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) u_\theta^s(a) - l_k(a) \right] + b_k \right\} \text{ s.t.}$$

$$\sum_{a \in \mathcal{A}} \left[\sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) u_\theta^k(a) + l_k(a) \right] - b_k \geq \sum_{a \in \mathcal{A}} y_{k,k',a} - b_{k'} \quad \forall k, k' \in \mathcal{K} : b_{k'} \leq b_k \tag{1a}$$

$$y_{k,k',a} \geq \sum_{\theta \in \Theta} \mu_\theta \phi_\theta^{k'}(a) u_\theta^k(a) + l_{k'}(a) \quad \forall k \in \mathcal{K}, \forall k' \in \mathcal{K}, \forall a \in \mathcal{A} \tag{1b}$$

$$y_{k,k',a} \geq \sum_{\theta \in \Theta} \mu_\theta \phi_\theta^{k'}(a) u_\theta^k(a') \quad \forall k \in \mathcal{K}, \forall k' \in \mathcal{K}, \forall a \neq a' \in \mathcal{A} \tag{1c}$$

$$\sum_{a \in \mathcal{A}} \left[\sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) u_\theta^k(a) + l_k(a) \right] - b_k \geq \sum_{\theta \in \Theta} \mu_\theta u_\theta^k(a') \quad \forall k \in \mathcal{K}, \forall a' \in \mathcal{A} \tag{1d}$$

$$\sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) u_\theta^k(a) + l_k(a) \geq \sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) u_\theta^k(a') \quad \forall k \in \mathcal{K}, \forall a \neq a' \in \mathcal{A} \tag{1e}$$

$$\sum_{a \in \mathcal{A}} \phi_\theta^k(a) = 1 \quad \forall k \in \mathcal{K}, \forall \theta \in \Theta. \tag{1f}$$

Eqs. (1a)–(1c) ensure IC, Eq. (1d) enforces IR, while Eq. (1e) guarantees persuasiveness. Given how LP (1) is obtained, it is *not* immediately clear how, given a feasible solution to LP (1), one can recover an IC and IR protocol with seller's expected utility equal to the value of the objective of LP (1). Indeed, in a solution to LP (1), it could be $l_k(a) > 0$ even when the variables $\phi_\theta^k(a)$ are equal to zero. Then, it is *not* possible to immediately recover $\pi_k(a)$ starting from a solution to LP (1), since $l_k(a)$ encodes $\sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) \pi_k(a)$, and computing $\pi_k(a)$ would require a division by zero. The following two results show that it is possible to efficiently recover an optimal protocol with menus starting from an optimal solution to LP (1).

Lemma 3.3. *The optimal value of LP* (1) *is at least as large as the supremum of the seller's expected utility over all the IC and IR protocols with menus.*

Lemma 3.4. *Given a feasible solution to LP* (1)*, one can recover in poly-time an IC and IR protocol with seller's expected utility at least the value of the solution to LP* (1)*.*

Lemma 3.4 holds since, given a feasible solution to LP (1), one can efficiently construct a new solution in which, if $l_k(a) > 0$, then $\sum_{\theta \in \Theta} \mu_\theta \phi_\theta^k(a) > 0$. Moreover, such a procedure does *not*

[10]Here we focus on protocols in which each action recommended by the seller maximizes the seller's expected utility among the buyer's best responses. If multiple best responses yield the same maximal expected utility for the seller, the buyer follows the recommendation.

decrease the objective function value and retains the IC and IR conditions. Then, from the newly obtained solution, one can easily recover the desired protocol by letting $\pi_k(a) = \frac{l_k(a)}{\sum_{\theta\in\Theta}\mu_\theta\phi^k_\theta(a)}$ when the denominator is positive, and $\pi_k(a) = 0$ otherwise, for all $k \in \mathcal{K}$ and $a \in \mathcal{A}$. Finally, we can design a polynomial-time algorithm that finds an optimal protocol with menus. Indeed, the algorithm simply solves LP (1) and uses Lemma 3.4 to recover an IC and IR protocol having at least the same value. Thanks to Lemma 3.3, such a protocol is optimal for the seller.

Theorem 3.1. *There exists a poly-time algorithm that computes an optimal protocol with menus.*

The idea of formulating a linear relaxation of a quadratic problem by introducing new variables has already been used in generalized principal-agent problems [19]. However, in such a setting, the relaxation cannot be used to solve the principal's optimization problem exactly, but only to recover a desirable approximation. This is because the problem may *not* admit a maximum [7]. Surprisingly, in our information-selling setting, the linear relaxation can be used to find (exactly) an optimal protocol. Intuitively, this is possible since, in our setting, the seller observes the buyer's action.

4 Connection With Principal-Agent Problems

4.1 Principal-Agent Problems With Observable Actions

First, we formally define *observable-action principal-agent problems*. For ease of exposition, we reuse some of the notation already introduced in Section 2. The *agent* has a finite set $\mathcal{K}$ of possible *types*, and $k \in \mathcal{K}$ is drawn with probability λ_k according to a probability distribution $\lambda \in \Delta_{\mathcal{K}}$. Each type $k \in \mathcal{K}$ has a set $\mathcal{A}$ of *actions*, with each action having a type-dependent *cost* $c^k_a \in [0, 1]$. The *principal* has a *reward* $r_a \in [0, 1]$ for every agent's action $a \in \mathcal{A}$. Moreover, the principal can commit to a *contract*, which can be encoded by a function $\pi : \mathcal{A} \to \mathbb{R}_+$ defining a payment $\pi(a)$ from the principal to the agent for every action $a \in \mathcal{A}$. Given a contract, a type $k \in \mathcal{K}$ plays a *best response* $b^k_\pi \in \mathcal{A}$, defined as $b^k_\pi \in \arg\max_{a\in\mathcal{A}} \left\{\pi(a) - c^k_a\right\}$, where we assume that ties are broken in favor of the principal. Finally, the principal's goal is to commit to a contract maximizing their expected utility, *i.e.*, $\sum_{k\in\mathcal{K}} \lambda_k[r_{b^k_\pi} - \pi(b^k_\pi)]$. See Appendix D for further details.

4.2 Reduction From Selling Information

We prove that information selling in the case in which protocols are without menus and the buyer has limited liability is related to the problem of finding an optimal contract in observable-action principal-agent problems. Specifically, given a posterior $\xi \in \Delta_\Theta$, designing a payment function π that maximizes the seller's expected utility under ξ is equivalent to finding an optimal contract in a suitably-defined principal-agent problem with observable actions. Formally, for ease of presentation, we introduce the notion of payment function that is optimal for the seller in a given posterior:

Definition 4.1. *Given a posterior $\xi \in \Delta_\Theta$, a payment function $\pi : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ is* optimal *in ξ if:*

$$\pi \in \operatorname*{argmax}_{\pi'} \sum_{k\in\mathcal{K}} \lambda_k \left[\sum_{\theta\in\Theta} \xi_\theta\, u^s_\theta(b^k_{\xi,\pi'}) - \pi'(\xi, b^k_{\xi,\pi'})\right].$$

In the problem above, the price p of the signaling scheme ϕ does *not* appear in the seller's expected utility, as it is restricted to limited liability settings. Thus, $p = 0$. Similarly, we can safely assume that all the buyer's types satisfy IR. Then, we have:

Lemma 4.1. *Given $\xi \in \Delta_\Theta$, solving the problem in Definition 4.1 is equivalent to computing a contract maximizing the principal's expected utility in an instance of observable-action principal-agent problem such that, for every type $k \in \mathcal{K}$ and action $a \in \mathcal{A}$:*

$$c^k_a = \sum_{\theta\in\Theta} \xi_\theta \left[u^k_\theta(b^k_\xi) - u^k_\theta(a)\right] \quad \textit{and} \quad r_a = \sum_{\theta\in\Theta} \xi_\theta\, u^s_\theta(a).$$

Moreover, any observable-action principal-agent instance can be reduced in polynomial time to computing a seller-optimal no-menu protocol with limited liability and a single state of nature.

Lemma 4.1 implies that, in instances where the buyer has limited liability and there is only one state of nature, one can compute an optimal protocol without menus by finding an optimal contract in an instance of observable-action principal-agent problem defined as in the lemma. Thus:

Theorem 4.1. *Restricted to instances in which the buyer has limited liability and there is only one state of nature, computing a seller-optimal protocol without menus is equivalent to the problem of finding an optimal contract in general instances of the observable-action principal-agent problem.*

While the computational complexity of finding optimal contracts in hidden-action principal-agent problems is well understood (see, *e.g.*, [9]), to the best of our knowledge, there are no results for problems with observable actions. Formally, we have:

Theorem 4.2. *In observable-action principal-agent problems, the problem of computing a contract maximizing the principal's expected utility is* APX*-hard.*

Corollary 4.1. *The problem of computing an optimal protocol without menus is* APX*-hard, even when the buyer has limited liability and the number of states of nature d is fixed.*

When the buyer has limited liability and the number of states of nature is fixed, one can efficiently compute a *bi-criteria* approximation of the optimal seller's expected utility. This is similar to the bi-criteria guarantees of [9] for hidden-action principal-agent problems.

5 Protocols Without Menus: The Case of a Buyer With Limited Liability

We now study the problem of computing an optimal protocol without menus when the buyer has limited liability, *i.e.*, $b_k = 0$ for each $k \in \mathcal{K}$. In Section 5.1, we show how to circumvent the APX-hardness result established in Corollary 4.1, by fixing the number of buyer's actions. In Appendix G, we show how to circumvent the APX-hardness by fixing the number of states of nature. In this section, we can assume w.l.o.g. that $p = 0$ and compactly denote a protocol with a pair (γ, π).

5.1 Fixing the Number of Buyer's Actions

We show that, whenever the buyer has limited liability and the number of buyer's actions m is fixed, the problem admits an additive PTAS, *i.e.*, we can design a protocol whose seller's expected utility is arbitrarily close to that of an optimal one in polynomial time. To design our PTAS, we start by observing that, since $p = 0$, the IR constraints are always satisfied. Thus, we can formulate the problem as follows:[11]

$$\max \sum_{k \in \mathcal{K}} \lambda_k \sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \left[\sum_{\theta \in \Theta} \xi_\theta \, u^s_\theta(b^k_{\xi,\pi}) - \pi(\xi, b^k_{\xi,\pi}) \right] \text{ s.t. } \sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \, \xi_\theta = \mu_\theta \quad \forall \theta \in \Theta.$$

The problem above is defined for distributions over posteriors γ with arbitrary finite supports, whose cardinality is not fixed in advance. Thus, the crucial result that we need to design a PTAS is the possibility of focusing on finite sets of posteriors. To do so, we introduce a class of posteriors, which are called *q-uniform*.

Definition 5.1. *A posterior $\xi \in \Delta_\Theta$ is* q-uniform *if it can be obtained by averaging the elements of a multi-set defined by $q \in \mathbb{N}_{>0}$ canonical basis vectors of $\mathbb{R}^d$.*

We denote by $\Xi_q \subseteq \Delta_\Theta$ the set of the q-uniform posteriors. It is easy to check that $|\Xi_q| \leq \min\{d^q, (q+1)^d\}$. First, we show that, given any posterior $\xi^* \in \Delta_\Theta$, payment function π, and $\epsilon > 0$, there always exists a signaling scheme γ supported on Ξ_q that induces ξ^* on average and guarantees a seller's expected utility close to that in ξ^* (assuming the buyer plays an ϵ-best response).

Lemma 5.1. *Given $\epsilon > 0$, $\alpha \in (0, 1)$, $\xi^* \in \Delta_\Theta$, and a payment function π, there always exists $\gamma \in \Delta_{\Xi_q}$ with $q = \lceil 2 \log(2m/\alpha)/\epsilon^2 \rceil$ such that*

$$\sum_{k \in \mathcal{K}} \lambda_k \sum_{\xi \in \Xi_q} \gamma_\xi \left[\sum_{\theta \in \Theta} \xi_\theta \, u^s_\theta(b^{k,\epsilon}_{\xi,\pi'}) - \pi'(\xi, b^{k,\epsilon}_{\xi,\pi'}) \right] \geq \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi^*_\theta \, u^s_\theta(b^k_{\xi^*,\pi}) - \pi(\xi^*, b^k_{\xi^*,\pi}) \Big] - \alpha,$$

[11] Under limited liability, an optimum exists. Payments above the seller's expected action reward can be set to zero without reducing the seller's payoff. The resulting compact set of posterior-payment pairs and seller-favorable tie-breaking yield an optimal Bayes-plausible distribution with finite support.

where we let π' be a payment function that is optimal in every posterior $\xi \in \Xi_q$ when the buyer plays an ϵ-best response, i.e., π' solves the problem in Definition 4.1 for every $\xi \in \Xi_q$ with $b^k_{\xi,\pi}$ replaced by $b^{k,\epsilon}_{\xi,\pi}$. Furthermore, the signaling scheme γ satisfies $\sum_{\xi \in \Xi_q} \gamma_\xi \, \xi_\theta = \xi^*_\theta \ \forall \theta \in \Theta$.

Another component that we need for our PTAS is an algorithm that, given a q-uniform posterior, computes an optimal payment function in that posterior. By Theorem 4.1, such an algorithm has to compute an optimal contract in observable-action principal-agent problems. By Theorem 4.2, this problem is APX-hard. We show that, whenever the number of buyer's actions m is fixed, the hardness result can be circumvented, by providing a polynomial-time algorithm that solves the desired task.

Lemma 5.2. *If the buyer has limited liability and m is fixed, it is possible to compute in polynomial time, for each $\xi \in \Delta_\Theta$, a payment function $\pi(\xi, \cdot)$ that is optimal for ξ.*

The final piece we need is a way of returning to best responses, rather than using ϵ-best responses. Corollary D.2 in Appendix D allows us to do so by modifying the payment function in all induced posteriors, losing only $2\sqrt{\epsilon}$ of the seller's expected utility.

We now design our PTAS. By Lemma 5.1, we restrict attention to signaling schemes supported on Ξ_q. Moreover, running the poly-time algorithm of Lemma 5.2 on each posterior in Ξ_q yields a payment function that is optimal for all of them. By Corollary D.2, this yields a seller's expected utility close to that of a payment function that is optimal in every q-uniform posterior under ϵ-best responses, allowing us to apply Lemma 5.1. Therefore, our PTAS solves a modified version of the original *non*-finite problem, where we set $\text{supp}(\gamma) \subseteq \Xi_q$ and use as payment function the one obtained by applying Lemma 5.2 to each $\xi \in \Xi_q$. This procedure requires polynomial time when m and the accuracy parameters α and ϵ are fixed, since $|\Xi_q| \leq d^q$ and $q = \left\lceil 2\log(2m/\alpha)/\epsilon^2 \right\rceil$ by Lemma 5.1.

Theorem 5.1. *In instances where the buyer has limited liability and the number of buyer's actions m is fixed, computing an optimal protocol without menus admits an additive PTAS.*

We show that a similar approach can be employed to derive a quasi-polynomial-time algorithm providing a *bi-criteria* approximation of the seller's expected utility in an optimal protocol, even when m is arbitrary. Indeed, in our PTAS, the computation of an optimal payment function in a given q-uniform posterior can be done in polynomial time only when the number of actions is fixed. However, an approximately-optimal function can be computed in polynomial time by applying Corollary D.1 in Appendix D. Moreover, since $q = \left\lceil 8\log(2m/\alpha)/\epsilon^2 \right\rceil$ and $|\Xi_q| \leq d^q$, the enumeration over q-uniform posteriors can be performed in time quasi-polynomial in the instance size. This gives the following:

Theorem 5.2. *If the buyer has limited liability, there exists an algorithm that, given $\alpha, \epsilon > 0$ and $\rho \in (0, 1/2]$ as input, returns a protocol without menus achieving a seller's expected utility at least $\rho \, \text{OPT} - 2^{-\Omega(1/\rho)} - (\alpha + 2\sqrt{\epsilon})$, where OPT is the optimal seller's expected utility. Moreover, for fixed α, ϵ, ρ, the algorithm runs in time polynomial in $\mathcal{I}^{\log m}$—where $\mathcal{I}$ is the size of the instance—, and the seller's expected utility in the returned protocol is greater than or equal to $\text{OPTLIN} - (\alpha + 2\sqrt{\epsilon})$, where OPTLIN is the best expected utility achieved by a protocol parametrized by β as in Corollary D.1.*

The second part of the statement above will be useful to compute an approximate protocol without menus in general settings. Intuitively, it says that, even if our approximation algorithm only gives a *bi-criteria* approximation, the returned protocol achieves a seller's expected utility arbitrarily close to that achievable with payment functions that set payments as a given fraction of the seller's utility.

6 Protocols Without Menus: The Case of a Buyer Without Limited Liability

We focus on protocols without menus when the buyer has *no* limited liability. In Appendix H, we provide a negative result for general instances in which the buyer may *not* have limited liability. This strengthens Corollary 4.1: the problem is NP-hard to approximate within a constant factor even when the number of buyer's actions m is fixed. We then show how to circumvent this hardness result by providing a quasi-polynomial-time *bi-criteria* approximation algorithm. In Appendix J, we also design a polynomial-time exact algorithm when the number of buyer's types is fixed.

6.1 A Quasi-Poly-Time Bi-Criteria Approximation

To prove the approximation guarantees of our algorithm, we rely on Theorems 5.2 and G.1, and decompose the seller's expected utility in an optimal protocol without menus into the sum of three terms. Our algorithm computes three protocols without menus, each approximating one of these terms. Choosing the best protocol among the three provides the desired result. In particular, we have:

- The *first term* is related to the seller's expected utility collected from buyer's types that do not participate. Such a utility term can be trivially achieved by a protocol that charges no price, discloses no information, and never pays back the buyer.
- The *second term* is related to the best seller's expected utility that can be extracted from a buyer's action. This is related to the optimal seller's expected utility with limited liability, where the seller can influence the buyer's action by using a signaling scheme and providing final payments, without charging an upfront price. Thus, this term can be approximated using either the algorithm of Theorem 5.2 or that of Theorem G.1.
- The *third term* is related to the seller's expected utility obtained by the transfers between the seller and the buyer, which include the charged price and the final payment. Such a utility term can be approximated by using a protocol that reveals all the information to the buyer while charging a carefully-chosen price for that.

Formally, we prove the following main result:

Theorem 6.1. *For any $\alpha > 0$ and $\rho \in (0, 1/6]$, there is an algorithm that computes a protocol without menus with seller's expected utility at least $\rho \operatorname{OPT} - 2^{-\Omega(1/\rho)} - \alpha$, where OPT is the optimal seller's expected utility. For fixed α and ρ, the algorithm runs in time polynomial in $\mathcal{I}^{\log m}$—where $\mathcal{I}$ is the size of the instance—when using Theorem 5.2 as a subroutine, and polynomial in $\mathcal{I}^d$ when using Theorem G.1.*

Acknowledgments

This publication was funded by the EU Horizon project ELIAS (European Lighthouse of AI for Sustainability, No. 101120237).

References

[1] Paola Alimonti and Viggo Kann. Some apx-completeness results for cubic graphs. *Theoretical Computer Science*, 237:123–134, 04 2000. doi: 10.1016/S0304-3975(98)00158-3.

[2] Tal Alon, Paul Dütting, and Inbal Talgam-Cohen. Contracts with private cost per unit-of-effort. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 52–69, 2021.

[3] Tal Alon, Paul Dütting, Yingkai Li, and Inbal Talgam-Cohen. Bayesian analysis of linear contracts, 2022.

[4] Moshe Babaioff, Robert Kleinberg, and Renato Paes Leme. Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 92–109, 2012.

[5] Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The design and price of information. *American Economic Review*, 108(1):1–48, January 2018. doi: 10.1257/aer.20161079. URL `https://www.aeaweb.org/articles?id=10.1257/aer.20161079`.

[6] Dirk Bergemann, Yang Cai, Grigoris Velegkas, and Mingfei Zhao. Is selling complete information (approximately) optimal? In *Proceedings of the 23rd ACM Conference on Economics and Computation*, EC '22, page 608–663, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391504. doi: 10.1145/3490486.3538304. URL `https://doi.org/10.1145/3490486.3538304`.

[7] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Designing menus of contracts efficiently: The power of randomization. *CoRR*, abs/2202.10966, 2022. URL `https://arxiv.org/abs/2202.10966`.

[8] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Designing menus of contracts efficiently: The power of randomization. In *EC '22: The 23rd ACM Conference on Economics and Computation*, pages 705–735, 2022.

[9] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Bayesian agency: Linear versus tractable contracts. *Artificial Intelligence*, 307:103684, 2022.

[10] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Multi-agent contract design: How to commission multiple agents with individual outcomes. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 412–448, 2023.

[11] Junjie Chen, Minming Li, and Haifeng Xu. Selling data to a machine learner: Pricing via costly signaling. In *International Conference on Machine Learning*, pages 3336–3359. PMLR, 2022.

[12] Yiling Chen, Haifeng Xu, and Shuran Zheng. Selling information through consulting. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2412–2431. SIAM, 2020.

[13] Constantinos Daskalakis and Vasilis Syrgkanis. Learning in auctions: Regret is hard, envy is easy. *Games and Economic Behavior*, 2022.

[14] Shaddin Dughmi and Haifeng Xu. Algorithmic bayesian persuasion. *SIAM Journal on Computing*, 50(3):STOC16–68, 2019.

[15] Shaddin Dughmi, Rad Niazadeh, Alexandros Psomas, and S Matthew Weinberg. Persuasion and incentives through the lens of duality. In *International Conference on Web and Internet Economics*, pages 142–155. Springer, 2019.

[16] Paul Dütting, Tim Roughgarden, and Inbal Talgam-Cohen. Simple versus optimal contracts. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, pages 369–387, 2019.

[17] Paul Dutting, Tim Roughgarden, and Inbal Talgam-Cohen. The complexity of contracts. *SIAM Journal on Computing*, 50(1):211–254, 2021.

[18] Paul Dütting, Michal Feldman, Inbal Talgam-Cohen, et al. Algorithmic contract theory: A survey. *Foundations and Trends® in Theoretical Computer Science*, 16(3-4):211–412, 2024.

[19] Jiarui Gan, Minbiao Han, Jibang Wu, and Haifeng Xu. Optimal coordination in generalized principal-agent problems: A revisit and extensions. *arXiv preprint arXiv:2209.01146*, 2022.

[20] Guru Guruganesh, Jon Schneider, and Joshua R Wang. Contracts under moral hazard and adverse selection. In *EC '21: The 22nd ACM Conference on Economics and Computation*, pages 563–582, 2021.

[21] Venkatesan Guruswami and Prasad Raghavendra. Hardness of learning halfspaces with noise. *SIAM Journal on Computing*, 39(2):742–765, 2009. doi: 10.1137/070685798. URL `https://doi.org/10.1137/070685798`.

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

[23] Shuze Liu, Weiran Shen, and Haifeng Xu. Optimal pricing of information. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 2021.

A Related Works

The study of algorithmic ways of selling information to an imperfectly-informed buyer has received some attention in the past. Babaioff et al. [4] initiated the study by considering a buyer with an unlimited budget. They provide an exponentially-sized *linear program* (LP) for computing an optimal mechanism for selling information, and they efficiently solve it through the ellipsoid method. The main drawback of the approach presented by Babaioff et al. [4] is that an optimal mechanism may require a significant money transfer from the buyer to the seller and *viceversa*, in order to only achieve a small overall net transfer. Chen et al. [12] complement the results in [4] by studying the problem of selling information when both the buyer and the seller are budget-constrained. Moreover, they also consider a setting in which the buyer's budget is private, and the seller needs to elicit it in the mechanism. Chen et al. [12] show that the addition of budget constraints considerably simplifies the problem of computing an optimal mechanism, since it can be formulated as a polynomially-sized LP, and thus solved efficiently.

The problem of selling information has also been addressed by Bergemann et al. [5], who study the case of binary actions and states of nature, characterizing a revenue-maximizing mechanism in such a setting. Furthermore, Bergemann et al. [6] extend the analysis to the case in which there are more than two actions and binary states of nature. In contrast, Liu et al. [23] study a revenue-maximizing mechanism for selling information when the stochasticity of the state of nature only affects a subset of the actions of the decision maker.

The problem of selling information has also been studied within the machine learning community, where the information being sold is data needed to train models; see, *e.g.*, [11].

Our problem is also related to the *Bayesian persuasion* framework originally introduced by Kamenica and Gentzkow [22], where an informed sender wants to influence the behavior of a self-interested receiver via the strategic provision of information. Dughmi et al. [15] generalize the classical framework by considering the case in which there are monetary transfers between the sender and the receiver. Our information-selling setting in which the buyer has limited liability generalizes the model of Dughmi et al. [15] by also introducing buyer's types.

Finally, let us remark that our information-selling problem shares critical features with *Bayesian principal-agent problems* (see, *e.g.*, the recent survey by [18] and [9, 2, 3, 20, 8, 10]). Indeed, as we show in Section 4, the problem of computing a seller-optimal protocol generalizes particular principal-agent problems in which the agent's action is observable. Such a connection between the two settings is also demonstrated in terms of results. In particular, notice that Castiglioni et al. [9] design bi-criteria approximation algorithms whose guarantees are similar to those provided in this paper. Moreover, Gan et al. [19] show how to find optimal protocols in generalized principal-agent problems by using a linear relaxation of the principal's optimization problem, which is quadratic.

B Motivating Examples

In Sections B.1 and B.2, we describe two examples of real-world scenarios in which our information-selling problem finds application. In Section B.3, we also discuss why it is relevant to consider protocols *without* menus in such real-world scenarios. In Section B.4, we provide an example showing that if the seller can provide a final payment to the buyer, they can significantly increase their revenue compared to the case in which they cannot. In Section B.5, we provide an example showing that mechanisms providing the buyer with a single contract can be sub-optimal compared to mechanisms that also exploit the information contained in the state of nature.

B.1 Start-Up Acquisition

Consider a scenario in which a multinational corporation (buyer) is interested in acquiring a start-up (seller). How much the corporation values the start-up depends on the financial health of the latter (state of nature), which is unknown to the corporation. The start-up, which has complete access to its own financial data, can choose to disclose some information about that (*e.g.*, private documentation and financial reports) to the corporation, in order to induce them to conclude the acquisition. In such a scenario, it is reasonable for the start-up to ask the corporation to leave a deposit (the price of information in our protocol) before accessing sensible data about the company, independently of the

outcome of the acquisition negotiation. Then, in the case in which the corporation decides to acquire the start-up, the latter may provide the former with a partial or total refund (encoded by the final, action-dependent payment in our protocol).

B.2 Real Estate Market

Consider a scenario in which a real estate agent (seller) wants to sell a house to a potential buyer. The agent is aware of the condition of the house (state of nature), while the buyer is not. In such a scenario, the agent may ask the buyer to express their interest in purchasing the house by signing an agreement involving an anticipated deposit (the price of information in our protocol), a house appraisal (corresponding to information revelation through a signaling scheme), and a potential final refund if the buyer decides not to buy the house after the appraisal (the final, action-dependent payment in our protocol).

B.3 On Protocols Without Menus

While protocols with menus allow the seller to maximize their expected utility in our information-selling problems with buyer's types, they also have some limitations that prevent them from being operationalized in practice. Indeed, in the vast majority of the real-world scenarios of interest, the set of possible buyer's types is considerably large. For instance, in the real estate market, the set of types has to encompass all the possible combinations of preferences and budgets of a potential buyer. Thus, since an optimal menu may contain one entry for each buyer's type (see Section 3), the menu may be prohibitively large to be reasonably adopted in practice. Indeed, it is unreasonable to ask a buyer to compare a huge number of alternatives among each other. Thus, in most of the cases, protocols without menus are the only viable option.

B.4 Payback vs. No-Payback Mechanisms

In the following, we provide a simple example showing that the seller's utility, when they cannot pay the buyer back, can be arbitrarily worse. We consider an instance with two states of nature $\Theta = \{\theta_1, \theta_2\}$, a uniform prior distribution, and two buyer's actions $\mathcal{A} = \{a_1, a_2\}$. Moreover, we let $u^s_{\theta_1}(a_2) = u^s_{\theta_2}(a_2) = 1$ and $u^s_{\theta_1}(a_1) = u^s_{\theta_2}(a_1) = 0$. Furthermore, we consider a buyer characterized by a single type k_1, with $b_{k_1} = 0$, $u^{k_1}_{\theta_1}(a_2) = u^{k_1}_{\theta_2}(a_2) = 0$ and $u^{k_1}_{\theta_1}(a_1) = u^{k_1}_{\theta_2}(a_1) = 0.5$. If the seller may not pay back the buyer, they will consistently choose action a_1 for any signaling scheme ϕ^{k_1} the seller may commit to. This is because they gain an expected utility of 0.5 by playing a_1, compared to an expected utility of 0 by playing a_2, regardless of the state of nature. Consequently, the seller's expected utility remains at 0. On the other hand, if the seller can provide repayment to the buyer, there exists a mechanism that provides a seller's utility of 0.5. Specifically, if the seller discloses no information by sending a single signal s for each state of nature and commits to the following payment scheme: $\pi_{k_1}(s, a_2) = 0.5$ and $\pi_{k_1}(s, a_1) = 0$, then the buyer will always select action a_2, leading to a seller's expected utility of 0.5.

B.5 Single Contract vs. Our Protocol

We provide a simple example showing that our mechanism provides a better seller's utility with respect to a mechanism implementing a single contract, even in the case in which there is a single buyer's type. We consider an instance with two states of nature $\Theta = \{\theta_1, \theta_2\}$, a uniform prior distribution, and two buyer's actions $\mathcal{A} = \{a_1, a_2\}$. Moreover, we let $u^s_{\theta_1}(a_2) = u^s_{\theta_2}(a_2) = 0.8$ and $u^s_{\theta_1}(a_1) = u^s_{\theta_2}(a_1) = 0$. Furthermore, we consider a buyer characterized by a single type k_1, with $b_{k_1} = 0$, $u^{k_1}_{\theta_1}(a_2) = 1$, $u^{k_1}_{\theta_2}(a_2) = 0$ and $u^{k_1}_{\theta_1}(a_1) = u^{k_1}_{\theta_2}(a_1) = 1$. In the case in which the seller can only offer a contract to the buyer, the seller's utility is at most 0.3. In particular, an optimal contract should incentivize the buyer to play action a_2. The optimal contract that incentivizes this action is given by the following payment scheme: $\pi_{k_1}(a_1) = 0$ and $\pi_{k_1}(a_2) = 0.5$. In this way, the buyer always plays action a_2 and the seller's expected utility is equal to 0.3. Instead, in our setting there exists a mechanism that extracts a seller's utility of 0.4. In particular, the mechanism that fully reveals the state of nature and provides no payment achieves utility 0.4. When the state is θ_1 the buyer plays a_2 providing utility 0.8 to the seller, while when the state is θ_2 the buyer plays a_1 providing utility 0 to the seller.

C Proofs Omitted From Section 3

For the ease of presentation, we first extensively write the quadratic optimization problem encoding the task of computing an optimal protocol with menus.

$$\sup_{\substack{\phi_\theta^k(s)\geq 0\\ b_k\geq p_k\geq 0\\ \pi_k(s,a)\geq 0}} \sum_{k\in\mathcal{K}} \lambda_k \left[\sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta^k(s)\left[u_\theta^s(b^k_{\xi^s,\pi_k}) - \pi_k(s, b^k_{\xi^s,\pi_k})\right] + p_k\right] \quad \text{s.t.} \tag{2a}$$

$$\sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta^k(s)\left[u_\theta^k(b^k_{\xi^s,\pi_k}) + \pi_k(s, b^k_{\xi^s,\pi_k})\right] - p_k \geq \sum_{s\in\mathcal{S}} \max_{a\in\mathcal{A}} \sum_{\theta\in\Theta} \mu_\theta \phi_\theta^{k'}(s)\left[u_\theta^k(a) + \pi_{k'}(s,a)\right] - p_{k'} \quad \forall k,k'\in\mathcal{K}: p_{k'}\leq b_k \tag{2b}$$

$$\sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta^k(s)\left[u_\theta^k(b^k_{\xi^s,\pi_k}) + \pi_k(s, b^k_{\xi^s,\pi_k})\right] - p_k \geq \max_{a\in\mathcal{A}} \sum_{\theta\in\Theta} \mu_\theta u_\theta^k(a) \qquad \forall k\in\mathcal{K} \tag{2c}$$

$$\sum_{s\in\mathcal{S}} \phi_\theta^k(s) = 1 \qquad \forall k\in\mathcal{K}, \forall\theta\in\Theta. \tag{2d}$$

Lemma 3.1. *Given any IC and IR protocol, one can always recover an IC and IR protocol that is direct and persuasive, and it provides the seller with the same expected utility.*

Proof. We discard signals that occur with zero probability. Let $\{(\phi^k, p_k, \pi_k)\}_{k\in\mathcal{K}}$ be an IC and IR protocol such that there exist two signals $s_1, s_2 \in \mathcal{S}$ inducing the same best response $\bar{a}\in A$ for a given receiver's type $\bar{k}\in\mathcal{K}$, *i.e.*, $b^{\bar{k}}_{\xi^{s_1},\pi_{\bar{k}}} = b^{\bar{k}}_{\xi^{s_2},\pi_{\bar{k}}} = \bar{a}$.

In the following, we show how to replace $\phi^{\bar{k}}$ and $\pi_{\bar{k}}$ with a new signaling scheme $\bar{\phi}^{\bar{k}}$ and a new payment function $\bar{\pi}_{\bar{k}}$ by merging s_1, s_2 into a single signal $\bar{s}$. Formally we define: $\bar{\phi}^{\bar{k}}_\theta(\bar{s}) = \phi^{\bar{k}}_\theta(s_1) + \phi^{\bar{k}}_\theta(s_2)$ for each $\theta\in\Theta$. Similarly, we define:

$$\bar{\pi}_{\bar{k}}(\bar{s}, a) = \frac{\sum_{\theta\in\Theta}\mu_\theta\phi^{\bar{k}}_\theta(s_1)\pi_{\bar{k}}(s_1,a) + \sum_{\theta\in\Theta}\mu_\theta\phi^{\bar{k}}_\theta(s_2)\pi_{\bar{k}}(s_2,a)}{\sum_{\theta\in\Theta}\mu_\theta(\phi^{\bar{k}}_\theta(s_1) + \phi^{\bar{k}}_\theta(s_2))} \quad \forall a\in\mathcal{A}\,.$$

By linearity, $\bar{a}$ is a best response at $\bar{s}$, and any other best response at $\bar{s}$ is a best response at both s_1 and s_2. Since ties at s_1 and s_2 are broken in favor of the seller, $\bar{a}$ also maximizes the seller's expected utility among the buyer's best responses at $\bar{s}$.

Finally, we do not change all the other components of the protocol *i.e.*, we leave these components of $\{(\bar{\phi}^k, \bar{p}_k, \bar{\pi}_k)\}_{k\in\mathcal{K}}$ equal to the one in $\{(\phi^k, p_k, \pi_k)\}_{k\in\mathcal{K}}$. To prove the lemma we show that the protocol $\{(\bar{\phi}^k, \bar{p}_k, \bar{\pi}_k)\}_{k\in\mathcal{K}}$ achieves the same seller's expected utility of $\{(\phi^k, p_k, \pi_k)\}_{k\in\mathcal{K}}$, while satisfying the IC and IR constraints. As a first step, we observe that:

$$\sum_{s\in\mathcal{S}\setminus\{s_1,s_2\}}\sum_{\theta\in\Theta}\mu_\theta\bar{\phi}^{\bar{k}}_\theta(s)\left[u^{\bar{k}}_\theta(b^{\bar{k}}_{\xi^s,\bar{\pi}}) + \bar{\pi}_{\bar{k}}(s, b^{\bar{k}}_{\xi^s,\bar{\pi}})\right] + \sum_{\theta\in\Theta}\mu_\theta\bar{\phi}^{\bar{k}}_\theta(\bar{s})[u^{\bar{k}}_\theta(\bar{a}) + \bar{\pi}_{\bar{k}}(\bar{s},\bar{a})] - \bar{p}_{\bar{k}} =$$
$$\sum_{s\in\mathcal{S}\setminus\{s_1,s_2\}}\sum_{\theta\in\Theta}\mu_\theta\phi^{\bar{k}}_\theta(s)\left[u^{\bar{k}}_\theta(b^{\bar{k}}_{\xi^s,\pi}) + \pi_{\bar{k}}(s, b^{\bar{k}}_{\xi^s,\pi})\right] + \sum_{\theta\in\Theta}\mu_\theta\phi^{\bar{k}}_\theta(s_1)[u^{\bar{k}}_\theta(\bar{a}) + \pi_{\bar{k}}(s_1,\bar{a})] +$$
$$+ \sum_{\theta\in\Theta}\mu_\theta\phi^{\bar{k}}_\theta(s_2)[u^{\bar{k}}_\theta(\bar{a}) + \pi_{\bar{k}}(s_2,\bar{a})] - p_{\bar{k}}.$$

The latter equality holds by linearity and proves that the protocol $\{(\bar{\phi}^k, \bar{p}_k, \bar{\pi}_k)\}_{k\in\mathcal{K}}$ preserves the left-hand sides of the IR and IC constraints. Moreover, thanks to the convexity of the max operator, we can show that for every type $k\in\mathcal{K}$:

$$\max_{a\in\mathcal{A}}\sum_{\theta\in\Theta}\mu_\theta\phi^{\bar{k}}_\theta(s_1)[u^k_\theta(a) + \pi_{\bar{k}}(s_1,a)] + \max_{a\in\mathcal{A}}\sum_{\theta\in\Theta}\mu_\theta\phi^{\bar{k}}_\theta(s_2)[u^k_\theta(a) + \pi_{\bar{k}}(s_2,a)] - p_{\bar{k}} \geq$$
$$\max_{a\in\mathcal{A}}\sum_{\theta\in\Theta}\mu_\theta\bar{\phi}^{\bar{k}}_\theta(\bar{s})[u^k_\theta(a) + \bar{\pi}_{\bar{k}}(\bar{s},a)] - p_{\bar{k}}.$$

Then, by summing over the set $(\mathcal{S} \cup \{\bar{s}\}) \setminus \{s_1, s_2\}$, we notice that the value of the right-hand side of the IC constraints achieved by the protocol $\{(\bar{\phi}^k, \bar{p}_k, \bar{\pi}_k)\}_{k \in \mathcal{K}}$ is less or equal to the value achieved by $\{(\phi^k, p_k, \pi_k)\}_{k \in \mathcal{K}}$. Due to that, we can easily conclude that the new protocol preserves the IC and the IR constraints. Finally, by observing that the following equality holds:

$$\sum_{\theta \in \Theta} \mu_\theta \bar{\phi}^{\bar{k}}_\theta(\bar{s})[u^s_\theta(\bar{a}) - \bar{\pi}_{\bar{k}}(\bar{s}, \bar{a})] =$$
$$\sum_{\theta \in \Theta} \mu_\theta \phi^{\bar{k}}_\theta(s_1)[u^s_\theta(\bar{a}) - \pi_{\bar{k}}(s_1, \bar{a})] + \sum_{\theta \in \Theta} \mu_\theta \phi^{\bar{k}}_\theta(s_2)[u^s_\theta(\bar{a}) - \pi_{\bar{k}}(s_2, \bar{a})].$$

we can easily prove that the two protocols achieve the same seller's expected utility. Then, by iterating this procedure for each buyer's type and couple of signals until there are no two signals inducing the same best response for that type, we get a protocol that employs direct and persuasive signals. □

Lemma 3.2. *Given an IC, IR, direct, and persuasive protocol*[12]*, it is always possible to recover an IC, IR, direct, and persuasive protocol* $\{(\phi^k, \tilde{p}_k, \tilde{\pi}_k)\}_{k \in \mathcal{K}}$ *that provides the same seller's expected utility, and, for every type* $k \in \mathcal{K}$*, satisfies* $\tilde{p}_k = b_k$ *and* $\tilde{\pi}_k(a, a') = 0$ *for all* $a \neq a' \in \mathcal{A}$*.*

Proof. We first prove that by setting $\tilde{\pi}_k(a, a') = 0$ for each $a \neq a' \in A$ and $k \in \mathcal{K}$, the seller's expected utility does not change and the IC and IR constraints are preserved. First, we show that by taking $\tilde{\pi}_k(a, a') = 0$ for each $a \neq a' \in \mathcal{A}$ and $k \in \mathcal{K}$, the signaling schemes ϕ^k remain persuasive. Indeed, we have that:

$$\begin{aligned}\sum_{\theta \in \Theta} \mu_\theta \phi^k_\theta(a)\big[u^k_\theta(a) + \pi_k(a, a)\big] &\geq \sum_{\theta \in \Theta} \mu_\theta \phi^k_\theta(a)\big[u^k_\theta(a') + \pi_k(a, a')\big] \\ &\geq \sum_{\theta \in \Theta} \mu_\theta \phi^k_\theta(a) u^k_\theta(a').\end{aligned}$$

Moreover, the left-hand side of all the equations defining the IR and IC constraints does not change since it depends only from the payments $\tilde{\pi}_k(a, a) = \pi_k(a, a)$ for each $a' \in \mathcal{A}$ that remains unchanged. Furthermore, by setting $\tilde{\pi}_k(a, a') = 0$ for each $a \neq a' \in \mathcal{A}$ and $k \in \mathcal{K}$, the right-hand sides of the IC constraints achieve smaller or equal values, since, intuitively, we are setting to zero non-negative values. Hence, the IC constraints are satisfied. Finally, by observing that the left-hand sides of the IR constraints and the seller's expected utility do not embed terms $\tilde{\pi}_k(a, a')$ with $a \neq a' \in \mathcal{A}$ and $k \in \mathcal{K}$, we conclude the first part of the proof.

In the second part of the proof we show that it is always possible to define a protocol in which the seller asks to each buyer's type to deposit all their budget at the beginning of the interaction, achieving the same seller's expected utility and satisfying the constraints. Formally, we show that given a protocol $\{(\phi^k, p_k, \pi_k)\}_{k \in \mathcal{K}}$ the protocol $\{(\phi^k, \tilde{p}_k, \tilde{\pi}_k)\}_{k \in \mathcal{K}}$, with $\tilde{p}_k = b_k$ and $\tilde{\pi}_k(a) = \pi_k(a, a) + b_k - p_k$ for each $k \in \mathcal{K}$ and $a \in \mathcal{A}$, achieves the same seller's expected utility. Indeed by linearity we have:

$$\begin{aligned}&\sum_{k \in \mathcal{K}} \lambda_k \left[\sum_{a \in \mathcal{A}} \sum_{\theta \in \Theta} \mu_\theta \phi^k_\theta(a)\big[u^s_\theta(a) - \pi_k(a)\big] + p_k \right] \\ &\quad = \sum_{k \in \mathcal{K}} \lambda_k \left[\sum_{a \in \mathcal{A}} \sum_{\theta \in \Theta} \mu_\theta \phi^k_\theta(a)\big[u^s_\theta(a) - \tilde{\pi}_k(a)\big] + b_k \right].\end{aligned}$$

With similar arguments it is easy to check that the protocol $\{(\phi^k, \tilde{p}_k, \tilde{\pi}_k)\}_{k \in \mathcal{K}}$ satisfies the IC and IR constraints, concluding the lemma. □

Lemma 3.3. *The optimal value of LP* (1) *is at least as large as the supremum of the seller's expected utility over all the IC and IR protocols with menus.*

Proof. We show that for each protocol $\{(\phi^k, p_k, \pi_k)\}_{k \in \mathcal{K}}$ that is feasible for Problem (2), we can derive a solution to LP (1) with at least the same value. This is sufficient to prove the statement.

[12]Here we focus on protocols in which each action recommended by the seller maximizes the seller's expected utility among the buyer's best responses. If multiple best responses yield the same maximal expected utility for the seller, the buyer follows the recommendation.

By Lemma 3.1 and Lemma 3.2, we focus without loss of generality on protocols $\{(\phi^k, p_k, \pi_k)\}_{k\in\mathcal{K}}$ that are direct and persuasive. Then, we can build a solution $(\bar{\phi}, \bar{l}, \bar{y})$ to LP (1) letting

$$\bar{l}_k(a) = \sum_{\theta\in\Theta} \mu_\theta \phi_\theta^k(a)\pi_k(a)$$

for each $k \in \mathcal{K}$ and $a \in \mathcal{A}$, and

$$\bar{y}_{k,k',a} = \max\left\{\sum_{\theta\in\Theta} \mu_\theta \phi_\theta^{k'}(a)\left(u_\theta^k(a) + \pi_{k'}(a,a)\right), \max_{a'\neq a}\sum_{\theta\in\Theta}\mu_\theta\phi_\theta^{k'}(a)u_\theta^k(a')\right\}$$

for each $k, k' \in \mathcal{K}$ and $a \in \mathcal{A}$. Moreover, we let $\bar{\phi} = \phi$. It is easy to verify that the solution $(\bar{\phi}, \bar{l}, \bar{y})$ results feasible for LP (1). □

Lemma 3.4. *Given a feasible solution to LP* (1)*, one can recover in poly-time an IC and IR protocol with seller's expected utility at least the value of the solution to LP* (1)*.*

Proof. As a first step we show that from a feasible solution (ϕ, l, y) to LP (1) we can recover another solution with at least the same value in which if $l_k(a) > 0$ then $\sum_{\theta\in\Theta}\mu_\theta\phi_\theta^k(a) > 0$. Specifically, given a $k \in \mathcal{K}$ and an $a \in \mathcal{A}$ such that $\sum_{\theta\in\Theta}\mu_\theta\phi_\theta^k(a) = 0$ and $l_k(a) > 0$, let $\bar{a} \in \mathcal{A}$ satisfy $\sum_{\theta\in\Theta}\mu_\theta\phi_\theta^k(\bar{a}) > 0$. We now define a new feasible solution $(\bar{\phi}, \bar{l}, \bar{y})$ as follows:

- $\bar{l}_k(a) = 0$
- $\bar{l}_k(\bar{a}) = l_k(a) + l_k(\bar{a})$
- $\bar{y}_{k',k,a} = 0 \quad \forall k' \in \mathcal{K}$
- $\bar{y}_{k',k,\bar{a}} = y_{k',k,\bar{a}} + l_k(a) \quad \forall k' \in \mathcal{K},$

while we leave all the other terms variables equal to the ones in (ϕ, l, y). It is easy to check that the solution $(\bar{\phi}, \bar{l}, \bar{y})$achieves the same seller's expected utility while satisfying the constrains. Applying this procedure for each couple (k, a) such that $\sum_{\theta\in\Theta}\mu_\theta\phi_\theta^k(a) = 0$ and $l_k(a) > 0$, we obtain a new solution $(\tilde{\phi}, \tilde{l}, \tilde{y})$ such that if $\tilde{l}_k(a) > 0$ then $\sum_{\theta\in\Theta}\mu_\theta\tilde{\phi}_\theta^k(a) > 0$.

To recover a feasible protocol we just need to set payments as follows. For each couple (k, a), if $\sum_{\theta\in\Theta}\mu_\theta\tilde{\phi}_\theta^k(a) > 0$, we set $\tilde{\pi}_k(a) = \tilde{l}_k(a)/(\sum_{\theta\in\Theta}\mu_\theta\tilde{\phi}_\theta^k(a))$. Otherwise, we set $\tilde{\pi}_k(a) = 0$. Notice that the ratio $\tilde{l}_k(a)/(\sum_{\theta\in\Theta}\mu_\theta\tilde{\phi}_\theta^k(a))$ is always well defined. Moreover, it is easy to see that $\{\tilde{\phi}^k, \tilde{p}_k, \tilde{\pi}_k\}_{k\in\mathcal{K}}$, with $\tilde{p}_k = b_k$ for each $k \in \mathcal{K}$ is a feasible solution to Problem (2) with at least the same value as (ϕ, l, y) for LP (1) . This concludes the proof. □

Theorem 3.1. *There exists a poly-time algorithm that computes an optimal protocol with menus.*

Proof. The algorithm solves LP 1. By Lemma 3.3, this solution has value greater or equal to the supremum of Program 2. Then, exploiting Lemma 3.4, we can recover in polynomial-time a protocol with at least the same utility, *i.e.*, an optimal one. □

D Additional Preliminary Technical Results on Observable-Action Principal-Agent Problems

In this section, we recall two already-known results on hidden-action principal-agent problems. Clearly, these also hold for the specific case of observable-action principal-agent problems. By Theorem 4.1, such results can be easily cast to our information-selling problem. Indeed, we also show that one of them can be strengthened in our setting.

The first result that we are going to introduce makes use of linear contracts, which are payment schemes that pay the agent a given fraction of the principal's reward. Formally, in observable-action principal-agent problems, a contract $\pi : \mathcal{A} \to \mathbb{R}_+$ is said to be *linear* if there exists a $\beta \in [0, 1]$ such that $\pi(a) = \beta\, r_a$ for all $a \in \mathcal{A}$. Despite their simplicity, linear contracts provide good approximations with respect to general ones. In particular, the following holds:

Theorem D.1 (Essentially Theorem 3 by Castiglioni et al. [9])**.** *In an observable-action principal-agent problem, for any $\rho \in (0, 1/2]$, there exists a linear contract $\pi : \mathcal{A} \to \mathbb{R}_+$ such that:*

$$\sum_{k \in \mathcal{K}} \lambda_k \left[r_{b_\pi^k} - \pi(b_\pi^k) \right] \geq \rho \max_{\pi'} \sum_{k \in \mathcal{K}} \lambda_k \left[r_{b_{\pi'}^k} - \pi'(b_{\pi'}^k) \right] - 2^{-\Omega(1/\rho)}.$$

Moreover, such a linear contract is defined by a parameter $\beta = 1 - 2^{-i}$, for some $i \in \{1, \dots, \lfloor 1/2\rho \rfloor\}$.

We will make use of a stronger version of Theorem D.1, which applies to our setting and directly follows from the analysis of Castiglioni et al. [9] and Lemma 4.1. Formally:

Corollary D.1. *Given a posterior $\xi \in \Delta_\Theta$, for any $\rho \in (0, 1/2]$, there exists a payment function $\pi : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ such that $\pi(\xi, a) = \beta \sum_{\theta \in \Theta} \xi_\theta \, u_\theta^s(a)$ for every $a \in \mathcal{A}$, where $\beta \in [0, 1]$ is an (action-independent) parameter, and, additionally, the following holds:*

$$\sum_{k \in \mathcal{K}} \lambda_k \sum_{\theta \in \Theta} \xi_\theta \left[u_\theta^s(b_{\xi,\pi}^k) - \pi(\xi, b_{\xi,\pi}^k) \right] \geq \rho \sum_{k \in \mathcal{K}} \lambda_k \max_{a \in \mathcal{A}} \sum_{\theta \in \Theta} \xi_\theta \left[u_\theta^s(a) + u_\theta^k(a) - u_\theta^k(b_\xi^k) \right] - 2^{-\Omega(1/\rho)}.$$

Moreover, such a parameter β is equal to $1 - 2^{-i}$ for some $i \in \{1, \dots, \lfloor 1/2\rho \rfloor\}$.

Proof. As a first step, we notice that even if Castiglioni et al. [9] show that linear contracts provide the desired approximation with respect to optimal contracts, their proof can be extended to show that linear contracts provide the same approximation with respect to the optimal social welfare, *i.e.*, $\sum_{k \in \mathcal{K}} \lambda_k \max_{a \in \mathcal{A}} [r_a - c_a^k]$. To prove this result, it is sufficient to follow all the steps of Theorem 3 of [9] except for the one in which Observation 1 is employed to upperbound the value of the optimal contract with the social welfare. Finally, we can modify this result to hold in our setting exploiting Lemma 4.1. □

Finally, we recall a useful result that establishes a connection between agent's best responses and *approximate* best responses in principal-agent problems. Intuitively, such a result states that, given a contract under which the agent is allowed to play an ϵ-best response (for some $\epsilon \geq 0$), it is always possible to recover a new contract in which the agent must play an (exact) best response, by only incurring in a small loss in the principal's expected utility. Formally, given $\epsilon \geq 0$ and a contract $\pi : \mathcal{A} \to \mathbb{R}_+$, for every $k \in \mathcal{K}$, we let $\mathcal{B}_\pi^{k,\epsilon} \subseteq \mathcal{A}$ be the set of *ϵ-best-response* actions for an agent of type k. Such a set is made by all the actions $a \in \mathcal{A}$ such that $\pi(a) - c_a^k \geq \max_{a' \in \mathcal{A}} \left\{ \pi(a') - c_{a'}^k \right\} - \epsilon$. We denote by $b_\pi^{k,\epsilon} \in \mathcal{B}_\pi^{k,\epsilon}$ an ϵ-best-response action that is actually played by an agent of type k, assuming that ties are broken in favor of the principal, as usual. Then:

Theorem D.2 (Essentially Proposition A.4 by Dutting et al. [17])**.** *Given $\epsilon \in [0, 1]$, an instance of an observable-action principal-agent problem and a contract $\pi : \mathcal{A} \to \mathbb{R}_+$, there exists a contract $\pi' : \mathcal{A} \to \mathbb{R}_+$ such that $\pi'(a) = (1 - \sqrt{\epsilon}) \, \pi(a) + \sqrt{\epsilon} \, r_a$ for every $a \in \mathcal{A}$, and the following holds:*

$$\sum_{k \in \mathcal{K}} \lambda_k \left[r_{b_{\pi'}^k} - \pi'(b_{\pi'}^k) \right] \geq \sum_{k \in \mathcal{K}} \lambda_k \left[r_{b_\pi^{k,\epsilon}} - \pi(b_\pi^{k,\epsilon}) \right] - 2\sqrt{\epsilon}.$$

Theorem D.2 can be easily cast to our setting by means of Lemma 4.1. Formally:

Corollary D.2. *Given $\epsilon \in [0, 1]$, a posterior $\xi \in \Delta_\Theta$, and a payment function $\pi : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$, there exists a payment function $\pi' : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ such that $\pi'(\xi, a) = (1 - \sqrt{\epsilon}) \, \pi(\xi, a) + \sqrt{\epsilon} \sum_{\theta \in \Theta} \xi_\theta u_\theta^s(a)$ for every $a \in \mathcal{A}$, and the following holds:*

$$\sum_{k \in \mathcal{K}} \lambda_k \left[\sum_{\theta \in \Theta} \xi_\theta \, u_\theta^s(b_{\xi,\pi'}^k) - \pi'(\xi, b_{\xi,\pi'}^k) \right] \geq \sum_{k \in \mathcal{K}} \lambda_k \left[\sum_{\theta \in \Theta} \xi_\theta \, u_\theta^s(b_{\xi,\pi}^{k,\epsilon}) - \pi(\xi, b_{\xi,\pi}^{k,\epsilon}) \right] - 2\sqrt{\epsilon}.$$

Corollary D.2 will be crucial to provide our results in the following sections.

E Proofs Omitted From Section 4

Lemma 4.1. *Given $\xi \in \Delta_\Theta$, solving the problem in Definition 4.1 is equivalent to computing a contract maximizing the principal's expected utility in an instance of observable-action principal-agent problem such that, for every type $k \in \mathcal{K}$ and action $a \in \mathcal{A}$:*

$$c_a^k = \sum_{\theta \in \Theta} \xi_\theta \left[u_\theta^k(b_\xi^k) - u_\theta^k(a) \right] \quad \textit{and} \quad r_a = \sum_{\theta \in \Theta} \xi_\theta \, u_\theta^s(a).$$

Moreover, any observable-action principal-agent instance can be reduced in polynomial time to computing a seller-optimal no-menu protocol with limited liability and a single state of nature.

Proof. We start proving the first part of the statement. Given an instance of Problem (4.1), we build an instance of the observable-action principal-agent problem with $c_a^k = \sum_{\theta\in\Theta} \xi_\theta \left[u_\theta^k(b_\xi^k) - u_\theta^k(a)\right]$ and $r_a = \sum_{\theta\in\Theta} \xi_\theta\, u_\theta^s(a)$. To prove the equivalence between the two settings, we first show that the set of best-responses for the two problems coincides. Indeed, given a payment function π and a type $k \in \mathcal{K}$, let $a \in \mathcal{B}_{\xi,\pi}^k$. Then, for each $a' \in \mathcal{A}$

$$\begin{aligned}
\pi(a) - c_a^k &= \pi(a) - \sum_{\theta\in\Theta} \xi_\theta \left[u_\theta^k(b_\xi^k) - u_\theta^k(a)\right] \\
&= \pi(a) + \sum_{\theta\in\Theta} \xi_\theta u_\theta^k(a) - \sum_{\theta\in\Theta} \xi_\theta u_\theta^k(b_\xi^k) \\
&\geq \pi(a') + \sum_{\theta\in\Theta} \xi_\theta u_\theta^k(a') - \sum_{\theta\in\Theta} \xi_\theta u_\theta^k(b_\xi^k) \\
&= \pi(a') - c_{a'}^k,
\end{aligned}$$

showing that $a \in \mathcal{B}_\pi^k$. Similarly, we can prove that if $a \in \mathcal{B}_\pi^k$, then $a \in \mathcal{B}_{\xi,\pi}^k$ This implies that the set of best responses are equivalent for each payment function π, *i.e.*, $\mathcal{B}_\pi^k = \mathcal{B}_{\xi,\pi}^k$. Since the two problems also use the same tie-breaking rule, the selected best responses coincide. Hence, for every payment function π,

$$\sum_{k\in\mathcal{K}} \lambda_k \left[r_{b_\pi^k} - \pi(b_\pi^k)\right] = \sum_{k\in\mathcal{K}} \lambda_k \left[\sum_{\theta\in\Theta} \xi_\theta\, u_\theta^s(b_{\xi,\pi}^k) - \pi(\xi, b_{\xi,\pi}^k)\right],$$

showing the equivalence between the two problems. This proves the first part of the statement.

We now show that from an instance of the observable-action principal-agent problem we can always build an instance of the selling-information problem without menus and with only a single state of nature θ. In particular, we set $u_\theta^s(a) = r_a$ for each $a \in \mathcal{A}$, and $u_\theta^k(a) = 1 - c_a^k$ for each $a \in \mathcal{A}$ and $k \in \mathcal{K}$. Following an analysis similar to the first part of the proof, we can show that the two problems are equivalent. This concludes the proof. □

Theorem 4.2. *In observable-action principal-agent problems, the problem of computing a contract maximizing the principal's expected utility is* APX*-hard.*

Proof. We reduce from vertex cover in cubic graphs. Formally, it is NP-hard to approximate the size of the minimum vertex cover in cubic graphs with an approximation $(1+\varepsilon)$, for a given constant $\varepsilon > 0$ [1]. Let $\eta = \varepsilon/7$. We show that an $(1-\eta)$-approximation to the principal-agent problem with observable actions can be used to provide a $(1+\varepsilon)$ approximation to vertex cover, concluding the proof.

Consider an instance of vertex cover (V, E) with nodes V and edges E. Let $\rho = |V|$ and $\ell = |E|$. Given a vertex $v \in V$, we let $E(v)$ be the set of edges e such that v is one of the extreme of the edge e. Similarly, given an edge $e \in E$, let $V(e)$ be the set of vertexes v such that e is an edge with extreme v. We build an instance of the principal-agent problem with observable actions as follows. For each vertex $v \in V$, there exists an agent's type k_v, while for each $e \in E$, there exists a type k_e. For each vertex $v \in V$, there exists an action a_v and an additional action a_-. The cost of a type k_e, $e \in E$, is $c_{a_-}^{k_e} = 0$, $c_{a_v}^{k_e} = \frac{1}{2}$ if $e \in E(v)$ and 1 otherwise. The cost of a type k_v, $v \in V$, is $c_{a_v}^{k_v} = 0$, and 1 otherwise. Finally, the principal's utility is equal to 1 if the action is in $\{a_v\}_{v\in V}$, while is equal to 0 otherwise, *i.e.*, $u_{a_v}^s = 1$ for each $v \in V$ and $u_{a_-}^s = 0$. All the types are equally probable, *i.e.*, $\lambda_k = \frac{1}{\rho+\ell}$ for each $k \in \mathcal{K}$.

First, we show that if there exists a vertex cover $V^\star$ of size ν, the value of the problem is at least $\frac{(\rho-\nu)+\frac{\nu}{2}+\frac{\ell}{2}}{\rho+\ell}$. Consider the payment function such that $\pi(a_v) = \frac{1}{2}$ if $v \in V^\star$ and 0 otherwise. A type k_v with $v \notin V^\star$ plays the action a_v and receives a payment of 0. A type k_v with $v \in V^\star$ plays the action a_v and receives a payment of $\frac{1}{2}$. A type k_e plays an action a_v such that $e \in E(v)$ (this action

exists by construction) and receives a payment of $\frac{1}{2}$. It is easy to see that the expected seller's utility is $\frac{(\rho-\nu)+\frac{\nu}{2}+\frac{\ell}{2}}{\rho+\ell}$.

Suppose that there exists an algorithm that provides a $1-\eta$ approximation. This implies that the algorithm returns a solution, *i.e.*, a payment function π, with value at least $(1-\eta)\frac{\rho-\frac{\nu}{2}+\frac{\ell}{2}}{\rho+\ell}$. We show how to exploit the payment function π to build a vertex cover of size at most $(1+\varepsilon)\nu$ in polynomial time. In particular, given π we recover a vertex cover $\bar{V}$ of the desired size as follows. First, it is easy to see that we can set the payment $\pi(a_-)=0$ and payments $\pi(a_v)\in\{0,\frac{1}{2}\}$ for each $v\in V$ without decreasing the utility. Intuitively, payments are useful only to change the best response of a type k_e, $e\in E$, from a_- to a_v, $v\in V$. To do so, it is sufficient a payment of $\frac{1}{2}$. Then, let $\bar{E}$ be the set of edges $e\in E$ such that the best response of k_e is a_-, *i.e.*, $\bar{E}=\{e\in E: b_\pi^{k_e}=a_-\}$. Consider a edge $e\in\bar{E}$ and a vertex $v\in V(e)$. Since $e\in\bar{E}$ the payment $\pi(a_v)=0$ and no type k_e plays action a_v. Hence, if we modify the payments by letting $\pi(a_v)=\frac{1}{2}$ on the action a_v we have three effects: i) the type k_e changes the best response to a_v, ii) some other types $e'\in E$ could change from action a_- to a_v,[13] iii) the payment of type k_v increases by $\frac{1}{2}$. Overall the principal's total utility increases by $\frac{1}{2}\lambda_{k_e}$ since k_e changes from action a_- with payment 0 to action a_v with payment $\frac{1}{2}$, and it decreases by $-\frac{1}{2}\lambda_{k_v}$ as the payment to type k_v increases by $\frac{1}{2}$. Moreover, if other types $k_{e'}$, $e'\neq e$ change from a_- with payment 0 to action a_v with payment $\frac{1}{2}$ the principal's utility increases. This implies that the principal's utility does not decrease with this procedure. Hence, repeating this procedure we can build a payment function π with weakly greater utility such that all the agent's type plays actions a_v, $v\in V$. Then, let $\bar{V}$ be the set of vertexes with at least one agent of type k_e, $e\in E$ that plays this action, *i.e.*, $\bar{V}=\{v\in V: b_\pi^{k_e}=a_v, e\in E\}$. We show that $\bar{V}$ is an vertex cover of size at most $(1+\epsilon)\nu$, concluding the proof. Notice that we can set payment $\pi(a_v)=0$ for each $v\in V\setminus\bar{V}$ without decreasing the seller's utility. Removing the payment does not change any best response since the only type playing the action a_v is the type k_v, the utility $u^{k_v}(a_v)=1$, and the utility of playing any other action $a\neq a_v$ is $u^{k_v}(a)+\pi(a)\leq\frac{1}{2}$. Then, the principal's utility is

$$\frac{(\rho-|\bar{V}|)+\frac{1}{2}|\bar{V}|+\frac{1}{2}\ell}{\rho+\ell}=\frac{\rho-\frac{1}{2}|\bar{V}|+\frac{1}{2}\ell}{\rho+\ell},$$

since $\rho-|\bar{V}|$ types k_v, $v\in V$, play a_v and receive payment 0, $|\bar{V}|$ types k_v, $v\in V$, play a_v and receive payment $\frac{1}{2}$, and all the types k_e, $e\in E$, play an action a_v, $v\in V$ and receive payment $\frac{1}{2}$. Then, since the principal's utility is by assumption $\frac{\rho-\frac{1}{2}|\bar{V}|+\frac{1}{2}\ell}{\rho+\ell}\geq(1-\eta)\frac{\rho-\frac{\nu}{2}+\frac{\ell}{2}}{\rho+\ell}$, it holds

$$|\bar{V}|\leq\eta(2\rho+\ell)+\nu\leq(1+7\eta)\nu=(1+\varepsilon)\nu,$$

where the second inequality comes from $\rho=\frac{2}{3}\ell$ and $\ell\leq 3\nu$. This concludes the proof. □

F Proofs Omitted From Section 5

Lemma 5.1. *Given $\epsilon>0$, $\alpha\in(0,1)$, $\xi^*\in\Delta_\Theta$, and a payment function π, there always exists $\gamma\in\Delta_{\Xi_q}$ with $q=\lceil 2\log(2m/\alpha)/\epsilon^2\rceil$ such that*

$$\sum_{k\in\mathcal{K}}\lambda_k\sum_{\xi\in\Xi_q}\gamma_\xi\left[\sum_{\theta\in\Theta}\xi_\theta\, u_\theta^s(b^{k,\epsilon}_{\xi,\pi'})-\pi'(\xi,b^{k,\epsilon}_{\xi,\pi'})\right]\geq\sum_{k\in\mathcal{K}}\lambda_k\Big[\sum_{\theta\in\Theta}\xi^*_\theta\, u_\theta^s(b^k_{\xi^*,\pi})-\pi(\xi^*,b^k_{\xi^*,\pi})\Big]-\alpha,$$

where we let π' be a payment function that is optimal in every posterior $\xi\in\Xi_q$ when the buyer plays an ϵ-best response, i.e., π' solves the problem in Definition 4.1 for every $\xi\in\Xi_q$ with $b^k_{\xi,\pi}$ replaced by $b^{k,\epsilon}_{\xi,\pi}$. Furthermore, the signaling scheme γ satisfies $\sum_{\xi\in\Xi_q}\gamma_\xi\,\xi_\theta=\xi^*_\theta\ \forall\theta\in\Theta$.

Proof. Let $\tilde{\xi}\in\Xi_q$ be the empirical mean of q *i.i.d.* samples drawn according to $\xi^*\in\Delta_\Theta$, where each $\theta\in\Theta$ has probability ξ^*_θ of being sampled. Therefore, $\tilde{\xi}\in\Xi_q$ is a random vector supported on q-uniform posteriors with expectation $\xi^*\in\Delta_\Theta$. Moreover, let $\gamma\in\Delta_{\Xi_q}$ be a probability distribution

[13] Notice that there could be other actions a'_v with $\pi(a_{v'})=\frac{1}{2}$ that provides the same utility of a_v for both the principal and the agent.

such as, for each $\xi \in \Xi_q$, it holds $\gamma_\xi := \Pr(\tilde{\xi} = \xi)$. We build a new payment function $\tilde{\pi}$ such that for each $\xi \in \Delta_\Theta$ and $a \in \mathcal{A}$, we have $\tilde{\pi}(\xi, a) = \pi(\xi^*, a)$. Fix a type $k \in \mathcal{K}$. Moreover, we let $\Xi^k_{q,\epsilon}$ be the set of posteriors such that $\xi \in \Xi^k_{q,\epsilon}$ if and only if for each $a \in \mathcal{A}$ it holds:

$$\left| \sum_{\theta \in \Theta} \left(\xi_\theta u^k_\theta(a) - \xi^*_\theta u^k_\theta(a) \right) \right| \leq \frac{\epsilon}{2}. \tag{3}$$

Then, for each $\xi \in \Xi^k_{q,\epsilon}$, we have that $\mathcal{B}^k_{\xi^*,\pi} \subseteq \mathcal{B}^{k,\epsilon}_{\xi,\tilde{\pi}}$. In particular, for any $a^* \in \mathcal{B}^k_{\xi^*,\pi}$, $\xi \in \Xi^k_{q,\epsilon}$ and $a \in \mathcal{A}$:

$$\begin{aligned}
\sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a^*) + \tilde{\pi}(\xi, a^*) &\geq \sum_{\theta \in \Theta} \xi^*_\theta u^k_\theta(a^*) + \tilde{\pi}(\xi^*, a^*) - \frac{\epsilon}{2} && \text{(By Eq. (3) and the definition of } \mathcal{B}^k_{\xi^*,\pi}) \\
&\geq \sum_{\theta \in \Theta} \xi^*_\theta u^k_\theta(a) + \tilde{\pi}(\xi^*, a) - \frac{\epsilon}{2} \\
&\geq \sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a) + \tilde{\pi}(\xi, a) - \epsilon && \text{(By Equation (3))}
\end{aligned}$$

Thus, $a^* \in \mathcal{B}^{k,\epsilon}_{\xi,\tilde{\pi}}$.

For each $a \in \mathcal{A}$, let $\tilde{t}^k_a := \sum_{\theta \in \Theta} \tilde{\xi}_\theta u^k_\theta(a) + \tilde{\pi}(\tilde{\xi}, a)$ and $t^k_a := \sum_{\theta \in \Theta} \xi^*_\theta u^k_\theta(a) + \tilde{\pi}(\xi^*, a)$. By the Hoeffding's inequality we have that, for each $a \in \mathcal{A}$,

$$\Pr(|\tilde{t}^k_a - \mathbb{E}[\tilde{t}^k_a]| \geq \frac{\epsilon}{2}) \leq 2e^{-2q(\epsilon/2)^2} \leq 2e^{-\log(2m/\alpha)} \leq \frac{\alpha}{m}. \tag{4}$$

Moreover, Equation (3) and the union bound yield the following:

$$\begin{aligned}
\sum_{\xi \in \Xi^k_{q,\epsilon}} \gamma_\xi &= \Pr(\tilde{\xi} \in \Xi^k_{q,\epsilon}) \\
&= \Pr(\bigcap_{a \in \mathcal{A}} \left| \tilde{t}^k_a - t^k_a \right| \leq \frac{\epsilon}{2}) \\
&\geq 1 - \sum_{a \in \mathcal{A}} \Pr(\left| \tilde{t}^k_a - t^k_a \right| \geq \frac{\epsilon}{2}) \\
&\geq 1 - \alpha. && \text{(By Equation (4))}
\end{aligned}$$

Define a payment function $\bar{\pi}$ by

$$\bar{\pi}(\xi, a) = \begin{cases} \tilde{\pi}(\xi, a), & \tilde{\pi}(\xi, a) \leq \sum_\theta \xi_\theta u^s_\theta(a), \\ 0, & \text{otherwise.} \end{cases}$$

Every action gives the seller a nonnegative utility under $\bar{\pi}$. Let $a^* = b^k_{\xi^*,\pi}$. For $\xi \in \Xi^k_{q,\epsilon}$, if $\sum_\theta \xi_\theta u^s_\theta(a^*) - \tilde{\pi}(\xi, a^*) \geq 0$, its payment is unchanged and the other payments can only decrease. Hence a^* remains an ϵ-best response under $\bar{\pi}$. If this difference is negative, the seller's utility under $\bar{\pi}$ is already at least that difference. Since the buyer selects the seller-favorable ϵ-best response, we obtain

$$\begin{aligned}
\sum_{\xi \in \Xi_q} \gamma_\xi \left[\sum_\theta \xi_\theta u^s_\theta(b^{k,\epsilon}_{\xi,\bar{\pi}}) - \bar{\pi}(\xi, b^{k,\epsilon}_{\xi,\bar{\pi}}) \right] &\geq \sum_{\xi \in \Xi^k_{q,\epsilon}} \gamma_\xi \left[\sum_\theta \xi_\theta u^s_\theta(a^*) - \pi(\xi^*, a^*) \right] \\
&\geq \sum_{\xi \in \Xi_q} \gamma_\xi \left[\sum_\theta \xi_\theta u^s_\theta(a^*) - \pi(\xi^*, a^*) \right] - \alpha \\
&= \sum_\theta \xi^*_\theta u^s_\theta(a^*) - \pi(\xi^*, a^*) - \alpha,
\end{aligned}$$

where the second inequality uses $\sum_\theta \xi_\theta u^s_\theta(a^*) - \pi(\xi^*, a^*) \leq 1$ and $\gamma(\Xi_q \setminus \Xi^k_{q,\epsilon}) \leq \alpha$. Summing over types with weights λ_k and using the optimality of π' in each posterior gives the desired utility bound.

Finally, by definition of γ, we have that, for each $\theta \in \Theta$:

$$\sum_{\xi \in \Xi_q} \gamma_\xi \xi_\theta = \xi^*_\theta.$$

This concludes the proof. □

Lemma 5.2. *If the buyer has limited liability and m is fixed, it is possible to compute in polynomial time, for each $\xi \in \Delta_\Theta$, a payment function $\pi(\xi, \cdot)$ that is optimal for ξ.*

Proof. Given a posterior $\xi \in \Delta_\Theta$ and a tuple $a \in \mathcal{A}^{|\mathcal{K}|}$ we let $\Pi_a \subseteq \mathbb{R}^m_+$ be the set of payment functions π such that for each $k \in \mathcal{K}$ it holds $a_k \in \mathcal{B}^k_{\xi,\pi}$. Given an $a \in \mathcal{A}^{|\mathcal{K}|}$, the problem of computing an optimal payment function restricted to payment functions in Π_a can be formulated as follows:

$$\begin{aligned}
&\min_{\pi} \sum_{k \in \mathcal{K}} \lambda_k \pi(\xi, a_k) \quad \text{s.t.} \\
&\sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a_k) + \pi(\xi, a_k) \geq \sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a') + \pi(\xi, a') \forall a' \in \mathcal{A}, k \in \mathcal{K} \\
&\pi(\xi, a') \geq 0 \qquad \forall a' \in \mathcal{A}\,.
\end{aligned}$$

We observe that, for each tuple $a \in \mathcal{A}^{|\mathcal{K}|}$, the vertexes of the regions $\Pi_a \subseteq \mathbb{R}^m_+$ are identified by m of the common $O(nm^2 + m)$ constraints:

$$\begin{aligned}
&\sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a') + \pi(\xi, a') \geq \sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a'') + \pi(\xi, a'') \forall a' \neq a'' \in \mathcal{A}, \forall k \in \mathcal{K} \\
&\pi(\xi, a') \geq 0 \qquad \forall a' \in \mathcal{A}.
\end{aligned}$$

Hence, the total number of vertexes defining all the regions Π_a, $a \in \mathcal{A}^{|\mathcal{K}|}$, is at most $\binom{nm^2+m}{m} = O((nm^2 + m)^m)$. Finally, since the objective function for the fixed action tuple is linear in Π_a for each tuple $a \in \mathcal{A}^{|\mathcal{K}|}$, given the optimal tuple of induced actions $a^* \in \mathcal{A}^{|\mathcal{K}|}$ the optimum for that tuple is attained in one of the vertexes of Π_{a^*}. Seller-favorable tie-breaking can only improve the utility at that vertex. Moreover, there are overall $\left|\bigcup_{a \in \mathcal{A}^{|\mathcal{K}|}} V(\Pi_a)\right| = O((nm^2 + m)^m)$ vertexes, where $V(\cdot)$ denotes the set of vertexes of the polyhedron. Hence, when m is fixed, it is possible to enumerate in polynomial time over all the vertexes in $\bigcup_{a \in \mathcal{A}^{|\mathcal{K}|}} V(\Pi_a)$ and compute the optimal payment function. □

Theorem 5.1. *In instances where the buyer has limited liability and the number of buyer's actions m is fixed, computing an optimal protocol without menus admits an additive PTAS.*

Proof. Given two arbitrary constants $\alpha \in (0, 1),\ \epsilon \in (0, 1]$ we let (γ, π) be an optimal protocol. We show that an $(\alpha + 2\sqrt{\epsilon})$-optimal protocol (γ^*, π^*) can be computed in polynomial time. As a first step we define a signaling scheme γ^* supported in Ξ_q as follows:

$$\gamma^*_{\tilde{\xi}} = \sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \gamma^\xi_{\tilde{\xi}} \quad \forall \tilde{\xi} \in \Xi_q,$$

where $\gamma^\xi \in \Delta_{\Xi_q}$ is the signaling scheme satisfying Lemma 5.1 with $q = \left\lceil 2\log(2m/\alpha)/\epsilon^2 \right\rceil$. First we observe that $\gamma^* \in \Delta_{\Xi_q}$ satisfies the consistency constraints, indeed we have:

$$\sum_{\tilde{\xi} \in \Xi_q} \gamma^*_{\tilde{\xi}} \tilde{\xi}_\theta = \sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \sum_{\tilde{\xi} \in \Xi_q} \gamma^\xi_{\tilde{\xi}} \tilde{\xi}_\theta = \sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \xi_\theta = \mu_\theta \quad \forall \theta \in \Theta.$$

Moreover, let $\pi^* : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ be the optimal payment function in each $\tilde{\xi} \in \Xi_q$. We show that the protocol $(\gamma^*, \pi^*, 0)$ is $(\alpha + 2\sqrt{\epsilon})$-optimal. Let $\pi'' : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ be the optimal payment function in each $\xi \in \Xi_q$ when the buyer is playing an ϵ-best response, *i.e.*,

$$\pi''(\xi, \cdot) \in \arg\max_{\tilde{\pi}(\xi,\cdot)} \sum_{k \in \mathcal{K}} \lambda_k [\sum_\theta \xi_\theta u^s_\theta(b^{k,\epsilon}_{\xi,\tilde{\pi}}) - \tilde{\pi}(\xi, b^{k,\epsilon}_{\xi,\tilde{\pi}})].$$

Moreover, let $\pi' : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ be the payment function such that $\pi'(\xi, a) = (1 - \sqrt{\epsilon})\pi''(\xi, a) + \sqrt{\epsilon} \sum_{\theta \in \Theta} \xi_\theta u^s_\theta(a)$ for each ξ and $a \in \mathcal{A}$. Then, by the optimality of π^*, Corollary D.2, and Lemma 5.1, respectively,

$$\begin{aligned}
&\sum_{k \in \mathcal{K}} \lambda_k \sum_{\tilde{\xi} \in \Xi_q} \gamma^*_{\tilde{\xi}} \left[\sum_\theta \tilde{\xi}_\theta u^s_\theta(b^k_{\tilde{\xi}, \pi^*}) - \pi^*(\tilde{\xi}, b^k_{\tilde{\xi}, \pi^*}) \right] \\
&\quad \geq \sum_{k \in \mathcal{K}} \lambda_k \sum_{\tilde{\xi} \in \Xi_q} \gamma^*_{\tilde{\xi}} \left[\sum_\theta \tilde{\xi}_\theta u^s_\theta(b^{k,\epsilon}_{\tilde{\xi}, \pi''}) - \pi''(\tilde{\xi}, b^{k,\epsilon}_{\tilde{\xi}, \pi''}) \right] - 2\sqrt{\epsilon} \\
&\quad \geq \sum_{k \in \mathcal{K}} \lambda_k \sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \left[\sum_\theta \xi_\theta u^s_\theta(b^k_{\xi, \pi}) - \pi(\xi, b^k_{\xi, \pi}) \right] - \alpha - 2\sqrt{\epsilon}.
\end{aligned}$$

Notice that the optimal payment $\pi^* : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ in each $\xi \in \Xi_q$ can be computed in polynomial time employing Lemma 5.2. Hence, to compute the optimal signaling scheme $\gamma^* \in \Delta_{\Xi_q}$ we can solve the following LP:

$$\begin{aligned}
\max_{\gamma \geq 0} \quad & \sum_{k \in \mathcal{K}} \lambda_k \sum_{\xi \in \Xi_q} \gamma_\xi \left[\sum_{\theta \in \Theta} \xi_\theta u^s_\theta(b^k_{\xi, \pi^*}) - \pi^*(\xi, b^k_{\xi, \pi^*}) \right] \\
\text{s.t.} \quad & \sum_{\xi \in \Xi_q} \gamma_\xi \xi_\theta = \mu_\theta \qquad \forall \theta \in \Theta.
\end{aligned}$$

Note that since $|\Xi_q| = O(d^q)$, all the payment function $\pi^* : \Xi_q \times \mathcal{A} \to \mathbb{R}_+$ can be precomputed in polynomial time. Moreover, the LP has polynomially many variables and constraints and can be solved efficiently. Finally, the solution returned by the LP is $\alpha + 2\sqrt{\epsilon}$-optimal. This concludes the proof. □

Theorem 5.2. *If the buyer has limited liability, there exists an algorithm that, given $\alpha, \epsilon > 0$ and $\rho \in (0, 1/2]$ as input, returns a protocol without menus achieving a seller's expected utility at least $\rho \, \text{OPT} - 2^{-\Omega(1/\rho)} - (\alpha + 2\sqrt{\epsilon})$, where OPT is the optimal seller's expected utility. Moreover, for fixed α, ϵ, ρ, the algorithm runs in time polynomial in $\mathcal{I}^{\log m}$—where $\mathcal{I}$ is the size of the instance—, and the seller's expected utility in the returned protocol is greater than or equal to $\text{OPTLIN} - (\alpha + 2\sqrt{\epsilon})$, where OPTLIN is the best expected utility achieved by a protocol parametrized by β as in Corollary D.1.*

Proof. The proof follows the same steps of Theorem 5.1. However, we compute the best linear payment function in each q-uniform posterior instead of using Lemma 5.2. Keeping β fixed when changing posteriors, the same argument gives the bound for OPTLIN. □

G Computing Protocols Without Menus: The Case of a Buyer With Limited Liability and a Fixed Number of States of Nature

In this section, we study the case in which the buyer has limited liability and the number of states of nature d is fixed. We prove that, in such a setting, it is possible to compute a *bi-criteria* approximation of an optimal protocol without menus similar to that in Theorem 5.2, but in polynomial time. Notice that such a result circumvents the APX-hardness one provided in Corollary 4.1, as the latter is based on a reduction working with instances with only one state of nature.

Similarly to Section 5.1, we first show that it is possible to employ signaling schemes supported on the set Ξ_q of q-uniform posteriors (for a suitably-defined $q \in \mathbb{N}_{>0}$), by only suffering an arbitrarily small, additive loss in terms of seller's expected utility. While the following result is similar to the one obtained in Lemma 5.1, it is based on different techniques and, in particular, on the stability of the optimal seller's expected utility under small changes in the buyer's posterior. Formally:

Lemma G.1. *Given any $\alpha > 0$, a posterior $\xi^* \in \Delta_\Theta$, and a payment function $\pi : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ that is optimal in every posterior $\xi \in \Xi_q$ with $q = \lceil 9d/\alpha^2 \rceil$, there exists a signaling scheme $\gamma \in \Delta_{\Xi_q}$:*

$$\sum_{k \in \mathcal{K}} \lambda_k \sum_{\xi \in \Xi_q} \gamma_\xi \left[\sum_{\theta \in \Theta} \xi_\theta \, u^s_\theta(b^k_{\xi, \pi}) - \pi(\xi, b^k_{\xi, \pi}) \right] \geq \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi^*_\theta u^s_\theta(b^k_{\xi^*, \pi}) - \pi(\xi^*, b^k_{\xi^*, \pi}) \Big] - \alpha,$$

Furthermore, the signaling scheme γ satisfies: $\sum_{\xi \in \Xi_q} \gamma_\xi \, \xi_\theta = \xi^*_\theta \ \forall \theta \in \Theta$.

Proof. We define a payment function $\pi' : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ as follows: $\pi'(\xi, a) = \pi(\xi^*, a)$ for each $a \in \mathcal{A}$ and $\xi \in \Delta_\Theta$. Furthermore, we define:

$$I_\alpha(\xi^*) = \left\{ \xi \in \Delta_\Theta : \|\xi - \xi^*\|_\infty \leq \frac{1}{q} \right\},$$

as the neighborhood of the given posterior $\xi^* \in \Delta_\Theta$ and $\Xi(\xi^*) = I_\alpha(\xi^*) \cap \Xi_q$ its intersection with the set Ξ_q. Notice that $\xi^* \in \text{co}(\Xi(\xi^*))$.[14] We may assume $\alpha < 1$, since otherwise the claimed utility bound follows from nonnegativity of the optimal seller's expected utility and $u^s_\theta(a) \leq 1$.

We show that for each $\xi \in I_\alpha(\xi^*)$, it holds $b^k_{\xi^*,\pi'} \in \mathcal{B}^{k,\epsilon}_{\xi,\pi'}$, where $\epsilon \coloneqq \alpha^2/9$. As a first step, for any buyer or seller utility function u, since $\sum_\theta (\xi_\theta - \xi^*_\theta) = 0$, we have that

$$\left| \sum_{\theta \in \Theta} (\xi_\theta - \xi^*_\theta) u_\theta(a) \right| \leq \tfrac{1}{2} \|\xi - \xi^*\|_1 \leq \tfrac{d}{2} \|\xi - \xi^*\|_\infty \leq \epsilon/2 \ \ \forall a \in \mathcal{A},$$

Moreover, by the definition of best response and the previous inequality, we have that:

$$\begin{aligned}
\sum_{\theta \in \Theta} \xi_\theta u^k_\theta(b^k_{\xi^*,\pi'}) + \pi'(\xi, b^k_{\xi^*,\pi'}) &\geq \sum_{\theta \in \Theta} \xi^*_\theta u^k_\theta(b^k_{\xi^*,\pi'}) + \pi'(\xi^*, b^k_{\xi^*,\pi'}) - \epsilon/2 \\
&\geq \sum_{\theta \in \Theta} \xi^*_\theta u^k_\theta(b^k_{\xi,\pi'}) + \pi'(\xi^*, b^k_{\xi,\pi'}) - \epsilon/2 \\
&\geq \sum_{\theta \in \Theta} \xi_\theta u^k_\theta(b^k_{\xi,\pi'}) + \pi'(\xi, b^k_{\xi,\pi'}) - \epsilon \\
&\geq \sum_{\theta \in \Theta} \xi_\theta u^k_\theta(a') + \pi'(\xi, a') - \epsilon,
\end{aligned}$$

for each $a' \in \mathcal{A}$. This shows that $b^k_{\xi^*,\pi'} \in \mathcal{B}^{k,\epsilon}_{\xi,\pi'}$.

Let $\pi^* : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ be the payment function prescribed by Corollary D.2. Then, for each $\xi \in \Xi(\xi^*)$, by optimality of π and Corollary D.2, we have that:

$$\begin{aligned}
&\sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi_\theta u^s_\theta(b^k_{\xi,\pi}) - \pi(\xi, b^k_{\xi,\pi}) \Big] \\
&\quad \geq \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi_\theta u^s_\theta(b^k_{\xi,\pi^*}) - \pi^*(\xi, b^k_{\xi,\pi^*}) \Big] \\
&\quad \geq \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi_\theta u^s_\theta(b^{k,\epsilon}_{\xi,\pi'}) - \pi'(\xi, b^{k,\epsilon}_{\xi,\pi'}) \Big] - 2\sqrt{\epsilon} \\
&\quad \geq \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi_\theta u^s_\theta(b^k_{\xi^*,\pi'}) - \pi'(\xi, b^k_{\xi^*,\pi'}) \Big] - 2\sqrt{\epsilon} \\
&\quad \geq \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi^*_\theta u^s_\theta(b^k_{\xi^*,\pi'}) - \pi'(\xi^*, b^k_{\xi^*,\pi'}) \Big] - 2\sqrt{\epsilon} - \epsilon \\
&\quad \geq \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi^*_\theta u^s_\theta(b^k_{\xi^*,\pi'}) - \pi'(\xi^*, b^k_{\xi^*,\pi'}) \Big] - 3\sqrt{\epsilon} \\
&\quad = \sum_{k \in \mathcal{K}} \lambda_k \Big[\sum_{\theta \in \Theta} \xi^*_\theta u^s_\theta(b^k_{\xi^*,\pi}) - \pi(\xi^*, b^k_{\xi^*,\pi}) \Big] - 3\sqrt{\epsilon}.
\end{aligned}$$

This shows that the expected seller's utility decreases of at most $3\sqrt{\epsilon}$ when we consider sufficiently close posteriors. Hence, by Caratheodory's theorem we can decompose ξ^* as follows:

$$\sum_{\xi' \in \Xi(\xi^*)} \gamma^{\xi^*}_{\xi'} \, \xi'_\theta = \xi^*_\theta \quad \forall \theta \in \Theta$$

[14] Given a finite set A we denote with $\text{co}(A)$ the set containing all the convex combination of elements in A.

with $\gamma^{\xi^*} \in \Delta_{\Xi(\xi^*)}$, where we recall that $\xi^* \in \text{co}(\Xi(\xi^*))$. We show now that such a decomposition decreases the expected seller's utility only by the desired amount. Formally, we have that:

$$\begin{aligned}
&\sum_{k\in\mathcal{K}} \lambda_k \sum_{\xi'\in\Xi(\xi^*)} \gamma^{\xi^*}_{\xi'} \Big(\sum_{\theta\in\Theta} \xi'_\theta u^s_\theta(b^k_{\xi',\pi}) - \pi(\xi', b^k_{\xi',\pi}) \Big) \\
&\quad \geq \sum_{\xi'\in\Xi(\xi^*)} \gamma^{\xi^*}_{\xi'} \Big(\sum_{k\in\mathcal{K}} \lambda_k \Big[\sum_{\theta\in\Theta} \xi^*_\theta u^s_\theta(b^k_{\xi^*,\pi}) - \pi(\xi^*, b^k_{\xi^*,\pi}) \Big] - 3\sqrt{\epsilon} \Big) \\
&\quad = \sum_{k\in\mathcal{K}} \lambda_k \Big[\sum_{\theta\in\Theta} \xi^*_\theta u^s_\theta(b^k_{\xi^*,\pi}) - \pi(\xi^*, b^k_{\xi^*,\pi}) \Big] - 3\sqrt{\epsilon}.
\end{aligned}$$

Since $3\sqrt{\epsilon} \leq \alpha$, this concludes the proof. □

Similarly to the case of a fixed number of actions, we employ Lemma G.1 to restrict the attention to signaling schemes (and thus payment functions) supported on Ξ_q. Moreover, in this case, we can apply Corollary D.1 in each q-uniform posterior in order to compute in polynomial time a payment function that provides a *bi-criteria* approximation of the optimal seller's expected utility in such a posterior. Finally, we design an algorithm that solves the finite-posterior LP described in Section 5.1, where we set $\text{supp}(\gamma) \subseteq \Xi_q$ in its objective and consistency constraints, and we take as payment function the one obtained by putting together those computed by means of Corollary D.1 for each $\xi \in \Xi_q$. Finally, the overall procedure requires polynomial time, since $|\Xi_q| \leq (q+1)^d$ and the number of states of nature d is fixed, and achieves a *bi-criteria* approximation of the seller's expected utility in an optimal protocol. Formally:

Theorem G.1. *Restricted to instances in which the buyer has limited liability and the number of states of nature d is fixed, there exists an algorithm that, given $\alpha > 0$ and $\rho \in (0, 1/2]$ as input, returns in time polynomial in the instance size, $1/\alpha$, and $1/\rho$ a protocol without menus achieving a seller's expected utility greater than or equal to $\rho\,\text{OPT} - 2^{-\Omega(1/\rho)} - \alpha$, where* OPT *is the seller's expected utility in an optimal protocol. Moreover, the seller's expected utility in the returned protocol is greater than or equal to $\rho\,\text{OPTLIN} - 2^{-\Omega(1/\rho)} - \alpha$ where* OPTLIN *is the best expected utility achieved by a protocol parametrized by β as in Corollary D.1.*

Proof. Given a constant $\alpha > 0$ we let (γ, π) be an optimal protocol. As a first step, we show that there exists a protocol (γ^*, π^*) achieving a seller's expected utility of at least $\mathsf{APX} \geq \rho\mathsf{OPT} - 2^{-\Omega(1/\rho)} - \alpha$, where OPT is the utility achieved with (γ, π). Moreover, the payment function π^* is a linear function with parameter $\beta \in \{1 - 2^{-i}\}_{i\in\{1,\dots,\lfloor 1/(2\rho)\rfloor\}}$. We may assume $\alpha < 1$, since otherwise a protocol with nonnegative utility satisfies both bounds. We define a signaling scheme γ^* supported in Ξ_q as follows:

$$\gamma^*_{\tilde{\xi}} = \sum_{\xi\in\text{supp}(\gamma)} \gamma_\xi \gamma^\xi_{\tilde{\xi}} \quad \forall \tilde{\xi} \in \Xi_q,$$

where $\gamma^\xi \in \Delta_{\Xi_q}$ is the signaling scheme constructed as in the proof of Lemma G.1 with $q = \lceil \frac{18d}{\alpha^2} \rceil$. First we observe that $\gamma^* \in \Delta_{\Xi_q}$ satisfies the consistency constraints, indeed we have:

$$\sum_{\tilde{\xi}\in\Xi_q} \gamma^*_{\tilde{\xi}} \tilde{\xi}_\theta = \sum_{\xi\in\text{supp}(\gamma)} \gamma_\xi \sum_{\tilde{\xi}\in\Xi_q} \gamma^\xi_{\tilde{\xi}} \tilde{\xi}_\theta = \sum_{\xi\in\text{supp}(\gamma)} \gamma_\xi \xi_\theta = \mu_\theta \quad \forall \theta \in \Theta.$$

Moreover, we can define as $\pi^* : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ as the payment function computed (in polynomial time) with Corollary D.1 in each q-uniform posterior.

Let π' be the optimal payment function in every posterior. We show that the protocol (γ^*, π^*) achieves the desired approximation. Formally:

$$\begin{aligned}
&\sum_{k\in\mathcal{K}} \lambda_k \sum_{\tilde{\xi}\in\Xi_q} \gamma^*_{\tilde{\xi}} \Big(\sum_{\theta\in\Theta} \tilde{\xi}_\theta u^s_\theta(b^k_{\tilde{\xi},\pi^*}) - \pi^*(\tilde{\xi}, b^k_{\tilde{\xi},\pi^*}) \Big) \\
&\quad \geq \rho \left(\sum_{k\in\mathcal{K}} \lambda_k \sum_{\tilde{\xi}\in\Xi_q} \gamma^*_{\tilde{\xi}} \Big(\sum_{\theta\in\Theta} \tilde{\xi}_\theta u^s_\theta(b^k_{\tilde{\xi},\pi'}) - \pi'(\tilde{\xi}, b^k_{\tilde{\xi},\pi'}) \Big) \right) - 2^{-\Omega(1/\rho)} \\
&\quad = \rho \sum_{\xi\in\text{supp}(\gamma)} \gamma_\xi \Big(\sum_{k\in\mathcal{K}} \lambda_k \sum_{\tilde{\xi}\in\Xi_q} \gamma^\xi_{\tilde{\xi}} \big(\sum_{\theta\in\Theta} \tilde{\xi}_\theta u^s_\theta(b^k_{\tilde{\xi},\pi'}) - \pi'(\tilde{\xi}, b^k_{\tilde{\xi},\pi'}) \big) \Big) - 2^{-\Omega(1/\rho)}
\end{aligned}$$

$$\geq \rho \sum_{k \in \mathcal{K}} \lambda_k \sum_{\xi \in \text{supp}(\gamma)} \gamma_\xi \Big(\sum_{\theta \in \Theta} \xi_\theta u^s_\theta(b^k_{\xi,\pi}) - \pi(\xi, b^k_{\xi,\pi}) \Big) - 2^{-\Omega(1/\rho)} - \alpha.$$

The first inequality follows from Corollary D.1, the equality from the definition of γ^*, and the last inequality from the proof of Lemma G.1, optimality of π', and $\rho \leq 1$. This implies that since (γ^*, π^*) is feasible for the following LP, it has value at least $\rho\text{OPT} - 2^{-\Omega(1/\rho)} - \alpha$.

$$\max_{\gamma \geq 0} \sum_{k \in \mathcal{K}} \lambda_k \sum_{\xi \in \Xi_q} \gamma_\xi \Big[\sum_{\theta \in \Theta} \xi_\theta u^s_\theta(b^k_{\xi,\pi^*}) - \pi^*(\xi, b^k_{\xi,\pi^*}) \Big] \quad \text{s.t.}$$

$$\sum_{\xi \in \Xi_q} \gamma_\xi \xi_\theta = \mu_\theta \quad \forall \theta \in \Theta.$$

To find the desired approximation it is sufficient to compute π^* in each q-uniform posterior and solve the LP. Note that since $|\Xi_q| = O(q^d)$, the computation of the payment function $\pi^* : \Delta_\Theta \times \mathcal{A} \to \mathbb{R}_+$ and the computation of the previous LP require polynomial time in the instance size, $1/\alpha$, and $1/\rho$.

Finally, to prove the second part of the statement it is sufficient to notice that $\text{OPTLIN} \leq \text{OPT}$. □

The above approach will be useful for deriving our results in the general case in which the buyer has *no* limited liability.

H Computing Protocols Without Menus: A Stronger Hardness Result for the General Case of a Buyer Without Limited Liability

In this section, we provide a negative result for general instances in which the buyer has *no* limited liability that is stronger than the one established in Corollary 4.1. In particular, the latter result states that the seller's optimization problem is APX-hard even in the special case in which the buyer has limited liability and there is only one state of nature, relying on a reduction employing instances with an arbitrary number of actions m. Indeed, for the specific case in which the buyer has limited liability and the number of actions m is fixed, Theorem 5.1 provides an additive PTAS. Next, we show that, in general instances where the buyer may *not* have limited liability, the problem does not admit a PTAS unless $\mathsf{P} = \mathsf{NP}$, even when the number of buyer's actions m is fixed. To prove such an hardness result, we employ a result by Guruswami and Raghavendra [21] (see Theorem H.1 below), which is about the following *promise problem* related to the satisfiability of a fraction of linear equations with rational coefficients and variables restricted to the hypercube.[15]

Definition H.1 (LINEQ-MA$(1-\zeta, \delta)$ by Guruswami and Raghavendra [21])**.** *For any two constants $\zeta, \delta \in \mathbb{R}$ satisfying $0 \leq \delta \leq 1 - \zeta \leq 1$,* LINEQ-MA$(1-\zeta, \delta)$ *is the following promise problem: Given a set of linear equations $Ax = c$ over variables $x \in \mathbb{Q}^{n_{\text{var}}}$, with coefficients $A \in \mathbb{Q}^{n_{\text{eq}} \times n_{\text{var}}}$ and $c \in \mathbb{Q}^{n_{\text{eq}}}$, distinguish between the following two cases:*

- *there exists a vector $\hat{x} \in \{0,1\}^{n_{\text{var}}}$ that satisfies at least a fraction $1-\zeta$ of the equations;*
- *every possible vector $x \in \mathbb{Q}^{n_{\text{var}}}$ satisfies less than a fraction δ of the equations.*

Theorem H.1 (Guruswami and Raghavendra [21])**.** *For all the constants $\zeta, \delta \in \mathbb{R}$ which satisfy $0 < \delta \leq 1 - \zeta < 1$, the problem* LINEQ-MA$(1-\zeta, \delta)$ *is* NP*-hard.*

Then, Theorem H.1 allows us to prove the following hardness result:

Theorem H.2. *The problem of computing a seller-optimal protocol without menus does not admit a PTAS unless $\mathsf{P} = \mathsf{NP}$, even when the number of buyer's actions m is fixed.*

Proof. We introduce a reduction from LINEQ-MA$(1-\zeta, \delta)$ to the design of the optimal protocol, showing that for ζ and δ small enough, the following holds:

- *Completeness*: If an instance of LINEQ-MA$(1-\zeta, \delta)$ admits a $1-\zeta$ fraction of satisfiable equations when variables are restricted to lie in the hypercube $\{0,1\}^{n_{\text{var}}}$, then there exists a protocol that provides to the seller's expected utility at least of η, where η will be defined in the following;

[15] In the definition in [21], the vector $\hat{x}$ can be non-binary. However, Guruswami and Raghavendra [21] use a binary vector $\hat{x}$ in their proof and, thus, their hardness result also holds for our definition.

- *Soundness*: If at most a δ fraction of the equations can be satisfied, then every protocol provides to the seller's expected utility at most $\eta - c$, where c is a *constant* defined in the following.

In the rest of the proof, given a vector of variables $\mathbf{x} \in \mathbb{Q}^{n_{\text{var}}}$, for $i \in [n_{\text{var}}]$, we denote with x_i the component corresponding to the i-th variable. Similarly, for $j \in [n_{\text{eq}}]$, c_j is the j-th component of the vector $\mathbf{c}$, whereas, for $i \in [n_{\text{var}}]$ and $j \in [n_{\text{eq}}]$, the (j, i)-entry of $\mathbf{A}$ is denoted by A_{ji}.

Reduction As a preliminary step, we normalize the coefficients by letting $\bar{\mathbf{A}} \coloneqq \frac{1}{\tau}\mathbf{A}$ and $\bar{\mathbf{c}} \coloneqq \frac{1}{\tau^2}\mathbf{c}$, where we let $\tau \coloneqq 2M \max\left\{\max_{i\in[n_{\text{var}}], j\in[n_{\text{eq}}]} A_{ji}, \max_{j\in[n_{\text{eq}}]} c_j, n_{\text{var}}^2\right\}$ and M will be defined in the following. It is easy to see that the normalization preserves the number of satisfiable equations. Formally, the number of satisfied equations of $\mathbf{Ax} = \mathbf{c}$ is equal to the number of satisfied equations of $\bar{\mathbf{A}}\bar{\mathbf{x}} = \bar{\mathbf{c}}$, where $\bar{\mathbf{x}} = \frac{1}{\tau}\mathbf{x}$. For every variable $i \in [n_{\text{var}}]$, we define a state of nature $\theta_i \in \Theta$. Moreover, we introduce three additional states $\theta_0, \theta_1, \theta_2 \in \Theta$. The prior distribution $\mu \in \text{int}(\Delta_\Theta)$ is defined in such a way that $\mu_{\theta_i} = \frac{1}{2n_{\text{var}}^2}$ for every $i \in [n_{\text{var}}]$, while $\mu_{\theta_0} = \frac{n_{\text{var}}-1}{2n_{\text{var}}}$, $\mu_{\theta_1} = \frac{1}{4}$, and $\mu_{\theta_2} = \frac{1}{4}$ (notice that $\sum_{\theta\in\Theta} \mu_\theta = 1$). We define four buyer's types $k_j^1, k_j^2, k_j^3, k_j^4 \in \mathcal{K}$ for each equation $j \in [n_{\text{eq}}]$, where the probability of observing each buyer's type is $\frac{1}{8n_{\text{eq}}}$. Moreover, we define an additional type $k^\star$. All the types $k \in \mathcal{K}$ have budget $b_k = \nu/2$, where ν will be defined in the following. The buyer has 9 actions available, namely $\mathcal{A} \coloneqq \{a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$. Then, we define the utilities of the players, where the utility is 0 when not specified. For each $k_j^1, j \in [n_{\text{eq}}]$, the utilities are:

- $u_{\theta_i}^{k_j^1}(a_0) = \frac{1}{2}$ for each $i \in [n_{\text{var}}]$,
- $u_{\theta_i}^{k_j^1}(a_1) = \frac{1}{2} - \bar{A}_{ji} + \bar{c}_j$ for each $i \in [n_{\text{var}}]$,
- $u_{\theta_i}^{k_j^1}(a_2) = \frac{1}{2} + \bar{A}_{ji} - \bar{c}_j$ for each $i \in [n_{\text{var}}]$
- $u_{\theta_0}^{k_j^1}(a_0) = \frac{1}{2}$,
- $u_{\theta_0}^{k_j^1}(a_1) = \frac{1}{2} + \bar{c}_j$,
- $u_{\theta_0}^{k_j^1}(a_2) = \frac{1}{2} - \bar{c}_j$.
- $u_{\theta_1}^{k_j^1}(a_3) = \frac{1}{2} + 2\nu$,

For each $k_j^2, j \in [n_{\text{eq}}]$, the utilities are:

- $u_{\theta_i}^{k_j^2}(a_0) = \frac{1}{2} - \bar{A}_{ji} + \bar{c}_j$ for each $i \in [n_{\text{var}}]$,
- $u_{\theta_i}^{k_j^2}(a_7) = \frac{1}{2}$ for each $i \in [n_{\text{var}}]$
- $u_{\theta_0}^{k_j^2}(a_0) = \frac{1}{2} + \bar{c}_j$,
- $u_{\theta_1}^{k_j^2}(a_3) = \frac{1}{2} + 2\nu$,

For each type $k_j^3, j \in [n_{\text{eq}}]$ the utilities are:

- $u_{\theta_i}^{k_j^3}(a_0) = \frac{1}{2} + \bar{A}_{ji} - \bar{c}_j$ for each $i \in [n_{\text{var}}]$,
- $u_{\theta_i}^{k_j^3}(a_7) = \frac{1}{2}$ for each $i \in [n_{\text{var}}]$

- $u_{\theta_0}^{k_j^3}(a_1) = \frac{1}{2} - \bar{c}_j$,
- $u_{\theta_1}^{k_j^3}(a_3) = \frac{1}{2} + 2\nu$,

For each type k_j^4, $j \in [n_{\text{eq}}]$ the utilities are equivalent to the one of type k_j^1 but with the following differences:

- $u_{\theta}^{k_j}(a_5) = \frac{1}{2}$ for each $\theta \in \Theta$,
- $u_{\theta_i}^{k_j}(a_7) = \frac{1}{2}$ for each $i \in [n_{\text{var}}]$,

Finally, the utilities of type $k^\star$ are:

- $u_{\theta_1}^{k^\star}(a_6) = 1$,
- $u_{\theta}^{k^\star}(a_1) = 1$ for each $\theta \in \Theta$.

Moreover, we let $u_\theta^k(a_8) = \frac{1}{2}$ for every $k \in K$ and $\theta \in \Theta$. Finally, the utility of the seller is:

- $u_\theta^{\mathsf{s}}(a_6) = \frac{1}{4}$ for each $\theta \in \Theta$,
- $u_\theta^{\mathsf{s}}(a_5) = 4\nu$ for each $\theta \in \Theta$,
- $u_\theta^{\mathsf{s}}(a_0) = \nu$ for each $\theta \in \Theta$,
- $u_\theta^{\mathsf{s}}(a_7) = 2\nu$ for each $\theta \in \Theta$.

We recall that the utility is 0 when not defined explicitly.

Completeness Suppose that there exists a vector $\hat{\mathbf{x}} \in \{0,1\}^{n_{\text{var}}}$ such that at least a fraction $1-\zeta$ of the equations in $\mathbf{A}\hat{\mathbf{x}} = \mathbf{c}$ are satisfied. Let $X^1 \subseteq [n_{\text{var}}]$ be the set of variables $i \in [n_{\text{var}}]$ with $\hat{x}_i = 1$, while $X^0 \coloneqq [n_{\text{var}}] \setminus X^1$. Given the definition of $\bar{\mathbf{A}}$ and $\bar{\mathbf{c}}$, there exists a vector $\bar{\mathbf{x}} \in \{0, \frac{1}{\tau}\}^{n_{\text{var}}}$ such that at least a fraction $1-\zeta$ of the equations in $\bar{\mathbf{A}}\bar{\mathbf{x}} = \bar{\mathbf{c}}$ are satisfied, and, additionally, $\bar{x}_i = \frac{1}{\tau}$ for all the variables in $i \in X^1$, while $\bar{x}_i = 0$ whenever $i \in X^0$. Let us consider an (indirect) signaling scheme $\phi : \Theta \to \Delta_{\mathcal{S}}$ where the set of signals is $\mathcal{S} \coloneqq \{s_1, s_2, s_3\}$. Let $q \coloneqq \frac{n_{\text{var}}(n_{\text{var}}-1)}{\tau - |X^1|}$. For each $i \in [n_{\text{var}}]$, let $\phi_{\theta_i}(s_1) = q$ and $\phi_{\theta_i}(s_2) = 1-q$ if $i \in X^1$, while $\phi_{\theta_i}(s_2) = 1$ otherwise. Moreover, let $\phi_{\theta_0}(s_1) = 1$, $\phi_{\theta_1}(s_3) = 1$ and $\phi_{\theta_2}(s_2) = 1$. Then, all the other probabilities $\phi_\theta(s)$ are set to 0. It is easy to see that the signaling scheme is feasible. Moreover, we set the price $p = \nu/2$. Finally, we set $\pi(s_3, a_6) = 2\nu$ and all the other payments $\pi(s,a) = 0$.

Now, we compute the expected seller's utility due of each type of buyer.

- The buyer of type $k^\star$ in the posterior ξ^{s_3} plays the action a_6 and gets utility $\sum_{\theta\in\theta} \xi_\theta^{s_3} u_\theta^{k^\star}(a_6) + \pi(s_3, a_6) = 1 + 2\nu$. Moreover, in the other posteriors ξ^{s_1} and ξ^{s_2} the seller's utility is at least 0. Finally, the protocol is IR for the buyer since the expected utility declining the protocol is 1 while accepting it is $-\pi/2 + 1 \cdot \frac{3}{4} + (1+2\nu)\frac{1}{4} = 1$. Hence, the expected principal utility when the buyer's type is $k^\star$ is at least $\sum_{s\in\mathcal{S}} \sum_{\theta\in\theta} \xi_\theta^s u_\theta^{k^\star}(b_{\xi^s,\pi}^{k^\star}) = \frac{1}{16}$.
- Consider a buyer k_j^1, $j \in [n_{\text{eq}}]$, such that the j-th equality is satisfied by the vector $\hat{\mathbf{x}}$. Now, let us take the buyer's posterior $\xi^{s_1} \in \Delta_\Theta$ induced by the signal s_1. Let $h \coloneqq \frac{\frac{q}{n_{\text{var}}^2}}{\sum_{i\in X^1} \frac{q}{n_{\text{var}}^2} + \frac{n_{\text{var}}-1}{n_{\text{var}}}}$. Then, using the definition of ξ^{s_1}, it is easy to check that $\xi_{\theta_i}^1 = h$ for every $i \in X^{s_1}$, $\xi_{\theta_i}^{s_1} = 0$ for every $i \in X^0$, while $\xi_{\theta_0}^{s_1} = \frac{\frac{n_{\text{var}}-1}{n_{\text{var}}}}{\sum_{i\in X^1} \frac{q}{n_{\text{var}}^2} + \frac{n_{\text{var}}-1}{n_{\text{var}}}} = 1 - h\left|X^1\right|$. The buyer of type $k_j \in \mathcal{K}$

experiences a utility of $\sum_{\theta\in\Theta}\xi_\theta^{s_1}u_\theta^{k_j}(a_0)=\frac{1}{2}$ by playing action a_0. Instead, the utility she gets by playing a_1 is defined as follows:

$$\sum_{\theta\in\Theta}\xi_\theta^{s_1}u_\theta^{k_j}(a_1)=\sum_{i\in X^1}h\left(\frac{1}{2}-\bar{A}_{ji}+\bar{c}_j\right)+\xi_{\theta_0}^1\left(\frac{1}{2}+\bar{c}_j\right)=$$
$$=h\left|X^1\right|\left(\frac{1}{2}+\bar{c}_j\right)-h\sum_{i\in X^1}\bar{A}_{ji}+\left(1-h\left|X^1\right|\right)\left(\frac{1}{2}+\bar{c}_j\right)=$$
$$=\frac{1}{2}+\bar{c}_j-h\sum_{i\in X^1}\bar{A}_{ji}=\frac{1}{2}+\bar{c}_j-\frac{1}{\tau}\sum_{i\in X^1}\bar{A}_{ji}=\frac{1}{2},$$

where the second to last equality holds since $h=\frac{1}{\tau}$ (by definition of h and q), while the last equality follows from the fact that the j-th equation is satisfied, and, thus, $\frac{1}{\tau}\sum_{i\in X^1}\bar{A}_{ji}=\bar{c}_j$ (recall that $\bar{x}_i=\frac{1}{\tau}$ for all $i\in X^1$). Using similar arguments, we can write $\sum_{\theta\in\Theta}\xi_\theta^{s_1}u_\theta^{k_j}(a_2)=\frac{1}{2}$. Moreover, all the other actions have utility 0. Hence, the buyer plays a_0 in the posterior ξ^{s_1}. In posterior ξ^{s_2} induced by signal s_2, the utility of each action different from a_8 is strictly smaller than $\frac{1}{2}$. Hence, the buyer will play a_8, while in posterior ξ^{s_3} induced by signal s_3, the utility of action a_3 is $\frac{1}{2}+2\nu$ and the buyer will play a_3. Hence the expected utility of the buyer is $\frac{3}{4}\frac{1}{2}+\frac{1}{4}(\frac{1}{2}+2\nu)-p=\frac{1}{2}$. Moreover, the protocol is IR for the buyer since if she declines the protocol the utility is $\frac{1}{2}$ while if she accepts the protocol the utility is $\frac{1}{2}$. Hence, when the buyer's type is k_j^1 the expected seller's utility is $\nu\mu_{\theta_0}+\nu/2$.

- Consider a buyer k_j^2 or k_j^3 , $j\in n_{\text{eq}}$ such that the j-th equality is satisfied. A similar argument as before shows that in posterior ξ^{s_1} the buyer's optimal action is a_7, while in posterior ξ^{s_2}, the optimal action is a_8. In posterior ξ^{s_3}, the optimal action is a_3. Hence, the expected buyer's utility is $\frac{3}{4}\frac{1}{2}+\frac{1}{4}(\frac{1}{2}+2\nu)-p=\frac{1}{2}$. Hence, the protocol is IR for the buyer and provides expected seller's utility at least $2\nu\mu_{\theta_0}+\nu/2$.
- Consider a buyer k_j^4, $j\in[n_{\text{eq}}]$ such that the j-th equality is satisfied. The buyer has an utility similar to k_j^1 and plays the same best responses. Hence, it is indifferent in participating or not participating to the protocol. We assume that they break ties in favor of the seller and does not accept. She plays action a_5 and the expected seller's utility is 2ν.

Since all the other buyer's types provide positive utility —it never happens that the expected payment from the seller to the buyer exceeds the payment from the buyer to the seller—, the expected seller's utility is at least

$$\eta=\frac{1}{32}+(1-\zeta)\frac{1}{8}(\mu_{\theta_0}\nu+\nu/2)+(1-\zeta)\frac{1}{4}(\mu_{\theta_0}2\nu+\nu/2)+(1-\zeta)\frac{1}{8}2\nu$$

Soundness As a first step, we upperbound the expected seller's utility from each type. It is easy to see that the maximum expected utility that the seller can extract from the buyer's type $k^\star$ is at most $\frac{1}{32}$. Moreover, the maximum expected utility that the seller can extract from a buyer of type k_j^1, $j\in[n_{\text{eq}}]$, is at most $\frac{1}{8}(\nu)$. The maximum expected utility that the seller can extract from a buyer of type k_j^2 or k_j^3, $j\in[n_{\text{eq}}]$ is at most $\frac{1}{4}\frac{3}{2}\nu$. Finally, the maximum expected utility that the seller can extract from a buyer of type k_j^4, $j\in[n_{\text{eq}}]$, is $\frac{1}{8}2\nu$.

Using the previous upperbounds, we can bound the component of the utility due to each set of types. For each constant $t<1$, there exist constants $c=c(t)$, $\zeta=\zeta(t)$ such that if the expected utility is greater than $\eta-c$ then the expected utility from types k_j^1, $j\in[n_{\text{eq}}]$, is at least $t\frac{1}{8}(\nu)$, the expected utility from types k_j^2, $j\in[n_{\text{eq}}]$, and k_j^3, $j\in[n_{\text{eq}}]$, is at least $t\frac{1}{4}\frac{3}{2}\nu$, and the expected utility from types k_j^4, $j\in[n_{\text{eq}}]$, is at least $t\frac{1}{8}2\nu$. To see that, consider for instance the types k_j^1, $j\in[n_{\text{eq}}]$. It must hold:

$$\frac{1}{32}+\bar{t}\frac{1}{8}\nu+\frac{1}{4}\frac{3}{2}\nu+\frac{1}{8}2\nu\geq\frac{1}{32}+(1-\zeta)\frac{1}{8}(\mu_{\theta_0}\nu+\nu/2)+(1-\zeta)\frac{1}{4}(\mu_{\theta_0}2\nu+\nu/2)+(1-\zeta)\frac{1}{8}2\nu$$

Since for n_{var} large enough μ_{θ_0} is close to $\frac{1}{2}$, for $c(t)$, $\zeta(t)$ small enough constant the equation is satisfied for $\bar{t}\geq t$. A similar result holds for every other set of types k_j^2 with $j\in[n_{\text{eq}}]$, k_j^3 with $j\in[n_{\text{eq}}]$, and k_j^4 with $j\in[n_{\text{eq}}]$.

The next step is to show the existence of a posterior in which a t fraction of agent of types k_j^1, $j \in [n_{\text{eq}}]$, play a_0 and the the same holds for each other set of types k_j^2,k_j^3 with action a_7. Suppose by contradiction that there is no posterior in which a t fraction of k_j^1, $j \in [n_{\text{eq}}]$, plays a_0. First, notice that the maximum payment is at most $p = \nu/2 + \frac{1}{M}$, otherwise all the buyer's types k_j^1 are not IR. Moreover, the seller's utility minus payment is greater than 0 in a posterior only if the agent plays a_0. Finally, it is easy to see that it is sufficient to consider signaling schemes that induce posteriors such that if $\xi_{\theta_i} > 0$, then $\xi_{\theta_1} = 0$ and $\xi_{\theta_2} = 0$ since states ξ_{θ_1} and ξ_{θ_2} disincentivize the actions with high seller's utility. Hence, the maximal utility from agents of types k_j^1 is at most

$$\frac{1}{8}\left[\nu/2 + \frac{1}{M} + (t - 1/n_{\text{eq}})\frac{1}{2}\nu\right] < t\frac{1}{8}\nu,$$

for M large enough, reaching a contradiction. A similar argument holds for the other types. This implies that there exists a set $Q \subseteq [n_{\text{eq}}]$ and a posterior ξ such that for each $j \in Q$ all the buyers k_j^1, k_j^2, and k_j^3 in the posterior play a_0,a_7, and a_7, respectively. Notice that $|Q| \geq 1 - 3(1-t)$ and for t large enough $|Q| > \delta$.

Suppose that there exists a signal inducing a posterior $\xi \in \Delta_\Theta$ in which all the buyer's types k_j^1, $j \in Q$ best respond by playing action a_0. We show that there exists at least one $j \in Q$ such that it holds $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_1) > \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_0)$ or $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_2) > \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_0)$. For every buyer's type $k_j^1 \in \mathcal{K}$, it holds $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j}(a_0) = \frac{1}{2}$. Moreover, it is the case that:

$$\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_1) = \sum_{i\in[n_{\text{var}}]} \xi_{\theta_i}\left(\frac{1}{2} - \bar{A}_{ji} + \bar{c}_j\right) + \xi_{\theta_0}\left(\frac{1}{2} + \bar{c}_j\right) = \frac{1}{2} + \bar{c}_j - \sum_{i\in[n_{\text{var}}]} \xi_{\theta_i}\bar{A}_{ji}.$$

Similarly, it holds:

$$\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_2) = \frac{1}{2} - \bar{c}_j + \sum_{i\in[n_{\text{var}}]} \xi_{\theta_i}\bar{A}_{ji}.$$

Suppose by contradiction that for every type k_j^1, $j \in Q$, it is the case that $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_0) \geq \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_1)$, which implies that $\bar{c}_j - \sum_{i\in[n_{\text{var}}]} \xi_{\theta_i}\bar{A}_{ji} \leq 0$, whereas it holds $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j^1}(a_0) \geq \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_j}(a_2)$, implying $-\bar{c}_j + \sum_{i\in[n_{\text{var}}]} \xi_{\theta_i}\bar{A}_{ji} \leq 0$. Thus, $\sum_{i\in[n_{\text{var}}]} \xi_{\theta_i}\bar{A}_{ji} = \bar{c}_j$ for every $j \in Q$ and the vector $\hat{\mathbf{x}} \in \mathbb{Q}^{n_{\text{var}}}$ with $\hat{x}_i = \xi_{\theta_i}$ for all $i \in [n_{\text{var}}]$ satisfies at least a fraction δ of the equations, reaching a contradiction. Since we have that t types k_j^1 play a_0, this implies that $\pi(\xi, a_0) > 0$. However, at the same time we have that the buyers of type k_j^2 and k_j^3 plays action a_7. Consider a $j^* \in Q$ such that $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^1}(a_1) > \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^1}(a_0)$ or $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^1}(a_2) > \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^1}(a_0)$. Recall that this buyer must play a_7. If the first inequality holds then it must hold $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^2}(a_7) + \pi(\xi, a_7) \geq \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^2}(a_0) + \pi(\xi, a_0)$. Moreover, $\sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^2}(a_7) = \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^1}(a_0) < \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^1}(a_1) = \sum_{\theta\in\Theta} \xi_\theta u_\theta^{k_{j^*}^2}(a_0)$, implying $\pi(\xi, a_7) > \pi(\xi, a_0)$. A similar argument holds for the buyer $k_{j^*}^3$ if the second inequality is satisfied. This implies that type $k_{j^*}^4$ can play the same best responses of player k_j^1 in any posterior different from ξ and play action a_7 in ξ. Hence, the expected utility of buyer $k_{j^*}^4$ is strictly greater than the one of $k_{j^*}^1$ (that is IR), and hence it is strictly IR.

We conclude the proof showing that the utility of this buyer's type is too small, reaching a contradiction. First, notice that the seller must induce a posterior with $\xi_{\theta_1} \geq \frac{3}{4}$ with probability at least $\frac{1}{8}$. In all the other posteriors the seller's utility from type k^* is 0. However, it must hold that the utility from type $k^\star$ is at least $\frac{1}{64}$ for ν small enough. Hence, playing posteriors with $\xi_{\theta_1} \geq \frac{3}{4}$ with probability smaller than $\frac{1}{8}$ the seller's utility form type k^* is at most $\frac{1}{2}\frac{1}{4}\frac{1}{8} < \frac{1}{64}$. Now consider the type $k_{j^*}^4$ that is IR. In a posterior ξ with $\xi_{\theta_1} \geq \frac{3}{4}$, the seller's utility when the type is $k_{j^*}^4$ is at most 0. Hence, the total utility from this type is at most $p + \frac{7}{8}4\nu \leq \nu + 1/M$, where the last inequality follows by the fact that the payment is at most $\frac{\nu}{2} + 1/M$. For $|Q|$ large enough, we have that a $|Q|/n_{\text{eq}} - \delta$ fraction of types k_j^4 provide seller's utility at most $\nu + 1/M$. Hence, the total utility

from type k_j^4 is at most $\frac{1}{8}[(|Q|/n_{\text{eq}} - \delta)(\nu + 1/M) + (1 - (|Q|/n_{\text{eq}} - \delta))2\nu] \leq t\frac{1}{8}2\nu$. Thus, we reach a contradiction. □

I Proofs Omitted From Section 6

Theorem 6.1. *For any $\alpha > 0$ and $\rho \in (0, 1/6]$, there is an algorithm that computes a protocol without menus with seller's expected utility at least $\rho \,\text{OPT} - 2^{-\Omega(1/\rho)} - \alpha$, where OPT is the optimal seller's expected utility. For fixed α and ρ, the algorithm runs in time polynomial in $\mathcal{I}^{\log m}$—where $\mathcal{I}$ is the size of the instance—when using Theorem 5.2 as a subroutine, and polynomial in $\mathcal{I}^d$ when using Theorem G.1.*

Proof. Let (ϕ, p, π) be an optimal protocol. Then, the seller's expected utility is given by:

$$\sum_{k \notin \mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_{\theta \in \Theta} \mu_\theta u_\theta^s(b_\mu^k) + \sum_{k \in \mathcal{R}_{\phi,p,\pi}} \lambda_k \left[\sum_{s \in \mathcal{S}} \sum_{\theta \in \Theta} \mu_\theta \phi_\theta(s) \left(u_\theta^s(b_{\xi^s,\pi}^k) - \pi(s, b_{\xi^s,\pi}^k) \right) + p \right],$$

where $\mathcal{R}_{\phi,p,\pi}$ is the set of buyer's types that participate under protocol (ϕ, p, π). Given a signal $s \in \mathcal{S}$ and a type $k \in \mathcal{K}$, let $b_{\xi^s}^k \in \arg\max_{a \in \mathcal{A}} \sum_{\theta \in \Theta} \mu_\theta \phi_\theta(s) u_\theta^k(a)$. Intuitively, $b_{\xi^s}^k$ is an optimal action for the buyer without considering the payment function. Then, the seller's utility can be spitted in three components:

(i) The utility from the buyer's types that do not participate

$$U_1 \coloneqq \sum_{k \notin \mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_{\theta \in \Theta} \mu_\theta u_\theta^s(b_\mu^k);$$

(ii) The seller's utility deriving from the buyer's action

$$U_2 \coloneqq \sum_{k \in \mathcal{R}_{\phi,p,\pi}} \lambda_k \left[\sum_{s \in \mathcal{S}} \sum_{\theta \in \Theta} \mu_\theta \phi_\theta(s) \left(u_\theta^s(b_{\xi^s,\pi}^k) + u_\theta^k(b_{\xi^s,\pi}^k) - u_\theta^k(b_{\xi^s}^k) \right) \right],$$

where we use the fact that to incentivize action $b_{\xi^s,\pi}^k$ over $b_{\xi^s}^k$ the payment must be at least $\sum_{\theta \in \Theta} \xi_\theta^s \left[u_\theta^k(b_{\xi^s}^k) - u_\theta^k(b_{\xi^s,\pi}^k) \right]$;

(iii) The utility related to the overall payment that the seller's can extract from the buyer given the price function π

$$U_3 \coloneqq \sum_{k \in \mathcal{R}_{\phi,p,\pi}} \lambda_k \left[p - \sum_s \sum_\theta \mu_\theta \phi_\theta(s) \left[\pi(s, b_{\xi^s,\pi}^k) + u_\theta^k(b_{\xi^s,\pi}^k) - u_\theta^k(b_{\xi^s}^k) \right] \right].$$

Notice that the term $U_2 + U_3$ is the utility deriving from participating buyer's types, where we add, respectively subtract, the term

$$\sum_{k \in \mathcal{R}_{\phi,p,\pi}} \lambda_k \left[\sum_{s \in \mathcal{S}} \sum_{\theta \in \Theta} \mu_\theta \phi_\theta(s) \left(u_\theta^k(b_{\xi^s,\pi}^k) - u_\theta^k(b_{\xi^s}^k) \right) \right]$$

to U_2, respectively U_3.

In the following, we design three protocols (ϕ^1, p^1, π^1), (ϕ^2, p^2, π^2), and (ϕ^3, p^3, π^3), each with seller's utility that approximates the corresponding utility terms U_1, U_2, and U_3. We will show that this will implies that at least one protocol provides a good approximation of the overall seller's utility, *i.e.*, of $U_1 + U_2 + U_3$.

Approximate U_1 The protocol (ϕ^1, p^1, π^1) that provides no information, charges no price, and does not provides any payment has seller's utility

$$\sum_{k \in \mathcal{K}} \lambda_k \sum_\theta \mu_\theta u_\theta^s(b_\mu^k) \geq \sum_{k \notin \mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_\theta \mu_\theta u_\theta^s(b_\mu^k) = U_1$$

Approximate U_2 By Corollary D.1, we know that for each signal $s \in \mathcal{S}$ (inducing a posterior ξ^s) and $\rho \in (0, 1/2]$, there exists a linear contract $\pi'(s,\cdot)$ such that $\pi'(s,a) = \beta \sum_{\theta\in\Theta} \xi^s_\theta u^s_\theta(a)$ with parameter $\beta = 1 - 2^{-i}$, $i \in \{1, \dots, \lfloor \frac{1}{2\rho} \rfloor\}$ that guarantees:

$$\sum_{k\in\mathcal{K}} \lambda_k \Big[\sum_{\theta\in\Theta} \xi^s_\theta \left(u^s_\theta(b^k_{\xi^s,\pi'}) - \pi'(\xi^s, b^k_{\xi^s,\pi'}) \right) \Big] \tag{7a}$$

$$\geq \rho \sum_{k\in\mathcal{K}} \lambda_k \max_{a\in\mathcal{A}} \left[\sum_\theta \xi^s_\theta \left(u^s_\theta(a) + u^k_\theta(a) - u^k_\theta(b^k_{\xi^s}) \right) \right] - 2^{-\Omega(1/\rho)} \tag{7b}$$

$$\geq \rho \sum_{k\in\mathcal{R}_{\phi,p,\pi}} \lambda_k \left[\sum_\theta \xi^s_\theta \left(u^s_\theta(b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi}) - u^k_\theta(b^k_{\xi^s}) \right) \right] - 2^{-\Omega(1/\rho)} \tag{7c}$$

where the first inequality comes from Corollary D.1, and the last one uses the nonnegativity of each maximum and evaluates the remaining maxima at $b^k_{\xi^s,\pi}$.

Now, we need a protocol (ϕ^2, p^2, π^2) that approximates U_2. When the number of states is fixed, we use the algorithm of Theorem G.1 and obtain a protocol with seller's expected utility at least

$$\rho U_2 - 2^{-\Omega(1/\rho)} - \alpha.$$

Otherwise, we can use Theorem 5.2 that is polynomial time when the number of actions is fixed, while it runs in quasi-polynomial time and provides a loss α when instantiated with sufficiently small parameters. Hence, protocol (ϕ^2, p^2, π^2) can be computed in time poly$(\min\{\mathcal{I}^d, \mathcal{I}^{log(m)}\})$. Notice that both the algorithms returns a protocol such that $p = 0$ and hence $p^2 = 0$. In the latter case, we can show that the protocol (ϕ^2, p^2, π^2) has seller's utility at least

$$\begin{aligned}
&\sum_{k\in\mathcal{K}} \lambda_k \left[\sum_{s\in\mathcal{S}} \sum_\theta \mu_\theta \phi_\theta(s) \left(u^s_\theta(b^k_{\xi^s,\pi'}) - \pi'(s, b^k_{\xi^s,\pi'}) \right) \right] - \alpha \\
&= \sum_{k\in\mathcal{K}} \lambda_k \left[\sum_{s\in\mathcal{S}} \left(\sum_\theta \mu_\theta \phi_\theta(s) \right) \sum_\theta \xi^s_\theta \left(u^s_\theta(b^k_{\xi^s,\pi'}) - \pi'(s, b^k_{\xi^s,\pi'}) \right) \right] - \alpha \\
&= \sum_{s\in\mathcal{S}} \left(\sum_\theta \mu_\theta \phi_\theta(s) \right) \sum_{k\in\mathcal{K}} \lambda_k \left[\sum_\theta \xi^s_\theta \left(u^s_\theta(b^k_{\xi^s,\pi'}) - \pi'(s, b^k_{\xi^s,\pi'}) \right) \right] - \alpha \\
&\geq \sum_{s\in\mathcal{S}} \left(\sum_\theta \mu_\theta \phi_\theta(s) \right) \left[\rho \sum_{k\in\mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_\theta \xi^s_\theta \left(u^s_\theta(b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi}) - u^k_\theta(b^k_{\xi^s}) \right) - 2^{-\Omega(1/\rho)} \right] - \alpha \\
&= \sum_{s\in\mathcal{S}} \left(\sum_\theta \mu_\theta \phi_\theta(s) \right) \left[\rho \sum_{k\in\mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_\theta \xi^s_\theta \left(u^s_\theta(b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi}) - u^k_\theta(b^k_{\xi^s}) \right) \right] - 2^{-\Omega(1/\rho)} - \alpha \\
&= \rho \sum_{k\in\mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_{s\in\mathcal{S}} \left(\sum_\theta \mu_\theta \phi_\theta(s) \right) \left[\sum_\theta \xi^s_\theta \left(u^s_\theta(b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi}) - u^k_\theta(b^k_{\xi^s}) \right) \right] - 2^{-\Omega(1/\rho)} - \alpha \\
&= \rho \sum_{k\in\mathcal{R}_{\phi,p,\pi}} \lambda_k \sum_{s\in\mathcal{S}} \sum_\theta \mu_\theta \phi_\theta(s) \left(u^s_\theta(b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi}) - u^k_\theta(b^k_{\xi^s}) \right) - 2^{-\Omega(1/\rho)} - \alpha,
\end{aligned}$$

where the initial bound holds since π' employs linear payments functions and (ϕ^2, p^2, π^2) has an additive loss α w.r.t. any protocol that employs linear payments functions, while the inequality in the chain comes from Equation (7).

Approximate U_3 Let $\delta_k \coloneqq \sum_\theta \mu_\theta u^k_\theta(b^k_\theta) - \sum_\theta \mu_\theta u^k_\theta(b^k_\mu)$ for each $k \in \mathcal{K}$, where b^k_θ is the best response of agent of type $k \in \mathcal{K}$ when the state of nature is θ. For each $k \in \mathcal{R}_{\phi,p,\pi}$, by the definition of IR it holds

$$\sum_{s\in\mathcal{S}} \sum_{\theta\in\Theta} \mu_\theta \phi_\theta(s) [\pi(s, b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi})] - p \geq \sum_\theta \mu_\theta u^k_\theta(b^k_\mu), \tag{8}$$

Hence,

$$p - \sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(s) \left[\pi(s, b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi}) - u^k_\theta(b^k_{\xi^s})\right] \tag{9a}$$

$$= p - \sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(s) \left[\pi(s, b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi})\right] + \sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(s) u^k_\theta(b^k_{\xi^s}) \tag{9b}$$

$$\leq p - \sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(s) \left[\pi(s, b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi})\right] + \sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(s) u^k_\theta(b^k_\theta) \tag{9c}$$

$$= p - \sum_{s\in\mathcal{S}}\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(s) \left[\pi(s, b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi})\right] + \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\theta) \tag{9d}$$

$$\leq - \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\mu) + \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\theta) \tag{9e}$$

$$\leq \delta_k, \tag{9f}$$

where the first inequality follows by the optimality of action b^k_θ in state θ, and the second one by Equation (8).

Moreover, the minimum-payment bound above implies that the left-hand side of Equation (9) is at most $p \leq b_k$, and hence at most $\min\{\delta_k, b_k\}$. Next, we show that for each $\zeta \in (0, 1]$ we can design a protocol with seller's utility of at least $\frac{\zeta}{2}\sum_{k\in\mathcal{K}} \lambda_k \min\{\delta_k, b_k\} - 2^{-\lfloor 1/\zeta \rfloor}$. Let $P_\zeta \coloneqq \{2^{-i}\}_{i\in\{1,\dots,\lfloor 1/\zeta\rfloor\}} \cup \{0\}$, and for each $k \in \mathcal{K}$ let p^k be the greatest $p \in P_\zeta$ such that $p \leq \min\{\delta_k, b_k\}$. Then,

$$\sum_{k\in\mathcal{K}} \lambda_k p^k \geq \sum_{k\in\mathcal{K}} \lambda_k \left(\min\{\delta_k, b_k\}/2 - 2^{-\lfloor 1/\zeta\rfloor}\right) = \sum_{k\in\mathcal{K}} \lambda_k \min\{\delta_k, b_k\}/2 - 2^{-\lfloor 1/\zeta\rfloor},$$

where the inequality holds since either $p^k \geq \min\{\delta_k, b_k\}/2$ or $\min\{\delta_k, b_k\} \leq 2^{-\lfloor 1/\zeta\rfloor}$

Hence, $\sum_{p\in P_\zeta} p \sum_{k\in\mathcal{K}: p^k = p} \lambda_k \geq \sum_{k\in\mathcal{K}} \lambda_k \min\{\delta_k, b_k\}/2 - 2^{-\lfloor 1/\zeta\rfloor}$, implying

$$\begin{aligned}\max_{p\in P_\zeta} p \sum_{k\in\mathcal{K}: p^k = p} \lambda_k &\geq \frac{1}{2(|P_\zeta| - 1)} \sum_{k\in\mathcal{K}} \lambda_k \min\{\delta_k, b_k\} - 2^{-\lfloor 1/\zeta\rfloor} \\ &\geq \frac{\zeta}{2} \sum_{k\in\mathcal{K}} \lambda_k \min\{\delta_k, b_k\} - 2^{-\lfloor 1/\zeta\rfloor}.\end{aligned}$$

Let $p^* = \operatorname{argmax}_{p\in P_\zeta} p \sum_{k\in\mathcal{K}: p^k = p} \lambda_k$. Consider the protocol (ϕ^3, p^3, π^3) that charges payment $p^3 = p^*$, reveals all information with ϕ^3 and set payment $\pi^3(s, a) = 0$ for each $s \in \mathcal{S}$ and $a \in \mathcal{A}$. We show that this protocol satisfies the IR constraint and is affordable for all the players such that $p^k = p^*$. Indeed, for all these types it holds

$$\begin{aligned}\sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\theta) - p^* &\geq \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\theta) - \delta_k \\ &= \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\theta) - \left(\sum_\theta \mu_\theta u^k_\theta(b^k_\theta) - \sum_\theta \mu_\theta u^k_\theta(b^k_\mu)\right) \\ &= \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\mu).\end{aligned} \tag{10}$$

For every type with $p^k = p^*$, the seller obtains at least p^*, whether the type participates or declines under seller-favorable tie-breaking. In particular, it is at least

$$\begin{aligned}p^* \sum_{k\in\mathcal{K}: p^k = p^*} \lambda_k &\geq \frac{\zeta}{2} \sum_{k\in\mathcal{K}} \lambda_k \min\{\delta_k, b_k\} - 2^{-\lfloor 1/\zeta\rfloor} \\ &\geq \frac{\zeta}{2} \sum_{k\in\mathcal{R}_{\phi,p,\pi}} \lambda_k \min\{\delta_k, b_k\} - 2^{-\lfloor 1/\zeta\rfloor}\end{aligned}$$

$$\geq \frac{\zeta}{2} \sum_{k \in \mathcal{R}_{\phi,p,\pi}} \lambda_k \left[p - \sum_{s \in \mathcal{S}} \sum_{\theta} \mu_\theta \phi_\theta(s) [\pi(s, b^k_{\xi^s,\pi}) + u^k_\theta(b^k_{\xi^s,\pi}) - u^k_\theta(b^k_{\xi^s})] \right] - 2^{-\lfloor 1/\zeta \rfloor}$$
$$= \frac{\zeta}{2} U_3 - 2^{-\lfloor 1/\zeta \rfloor},$$

where in the first inequality we use the bound on p^*, and in the third inequality we use Equation (9) and the budget bound above. Equivalently, setting $\rho = \zeta/2$, we obtain that for each $\rho \in (0, 1/2]$ there exists a protocol (ϕ^3, p^3, π^3) that has seller's utility at least $\rho U_3 - 2^{-\Omega(1/\rho)}$.

Wrapping up Let $i = \arg\max_{j \in \{1,2,3\}} U_j$ and OPT be the seller's utility with the optimal protocol (ϕ, p, π). Then, since $U_1 + U_2 + U_3 = \text{OPT}$, we have that $U_i \geq \frac{1}{3}\text{OPT}$. Moreover, since for each $\rho \in (0, 1/2]$ we can approximate each utility U_i, $i \in \{1, 2, 3\}$ with a protocol with utility at least $\rho U_i - 2^{-\Omega(1/\rho)} - \alpha$, the seller's utility of our approximation algorithm is at least $\rho U_i - 2^{-\Omega(1/\rho)} - \alpha \geq \rho\text{OPT}/3 - 2^{-\Omega(1/\rho)} - \alpha$. Finally, setting $\rho' = \rho/3$, we obtain that for each $\rho' \in (0, 1/6]$ the utility of the designed protocol is at least $\rho'\text{OPT} - 2^{-\Omega(1/\rho')} - \alpha$. This concludes the proof. □

J Computing Protocols Without Menus: The Case of a Buyer Without Limited Liability and a Fixed Number of Buyer's Types

In this section, we study the problem of computing a seller-optimal protocol without menus when the number of buyer's types is fixed, showing that it is possible to design a polynomial-time algorithm. As a byproduct, the existence of such an algorithm shows that, for protocols without menus, the seller's optimization problem always admits a maximum. As a preliminary result, we show that it is always possible to focus on protocols without menus that employ signals belonging to the set $\mathcal{A}^n$, and define signaling schemes $\phi : \Theta \to \Delta_{\mathcal{A}^n}$ and payment functions $\pi : \mathcal{A}^n \times \mathcal{A} \to \mathbb{R}_+$ such that, for every signal $\boldsymbol{a} \in \mathcal{A}^n$ and $k \in \mathcal{K}$, it holds $a_k \in \mathcal{B}^k_{\xi^{\boldsymbol{a}},\pi}$, where $a_k \in \mathcal{A}$ denotes the action corresponding to type k in $\boldsymbol{a}$. Intuitively, in such protocols, a signal specifies an action recommendation for each buyer's type, so that the buyer is always incentivized to follow such recommendations. With a slight abuse of notation, we say that protocols without menus (ϕ, p, π) as described above are *generalized-direct* and *generalized-persuasive*. Formally, we state the following lemma. To prove it, we observe that given a protocol, if two signals induce the same best response for every buyer type, they can always be merged while retaining the same expected utility for both the seller and the buyer. Then, by iterating such a process, we recover a signaling scheme and a payment function employing $\mathcal{A}^n$ as set of signals .

Lemma J.1. *Given a seller's protocol without menus, there always exists another protocol without menus that is generalized-direct and generalized-persuasive, achieving the same expected utility for the seller as the original protocol.*

Proof. Let (ϕ, π, p) be a protocol and let $s_1, s_2 \in \mathcal{S}$ be two signals such that $b^k_{\xi^{s_1},\pi} = b^k_{\xi^{s_2},\pi}$ for each receiver's type $k \in \mathcal{K}$. We show that it is always possible to define a new protocol (ϕ^*, π^*, p) that employs a single signal s^* instead of s_1 and s_2 achieving the same seller's expected utility while satisfying the constraints. Formally, we define a new signaling scheme ϕ^* as follows:

$$\begin{cases} \phi^*_\theta(s^*) = \phi_\theta(s_1) + \phi_\theta(s_2) \ \ \forall \theta \in \Theta \\ \phi^*_\theta(s) = \phi_\theta(s) \ \ \forall \theta \in \Theta, \ \ \forall s \in \mathcal{S} \setminus \{s_1, s_2\} \end{cases}$$

and a new payment function π^* as follows:

$$\begin{cases} \pi^*(s^*, a) = z\pi(s_1, a) + (1 - z)\pi(s_2, a) & \forall a \in \mathcal{A} \\ \pi^*(s, a) = \pi(s, a) \ \ \forall a \in \mathcal{A}, & \forall s \in \mathcal{S} \setminus \{s_1, s_2\} \end{cases}$$

As a first step, we observe that for each $k \in \mathcal{K}$ it holds:

$$\sum_{\theta \in \Theta} \mu_\theta \Big[\phi_\theta(s_1) \left[u^s_\theta(b^k_{\xi^{s_1},\pi}) - \pi(s_1, b^k_{\xi^{s_1},\pi}) \right] + \phi_\theta(s_2) \left[u^s_\theta(b^k_{\xi^{s_2},\pi}) - \pi(s_2, b^k_{\xi^{s_2},\pi}) \right] \Big] =$$

$$\sum_{\theta\in\Theta}\mu_\theta\phi^*_\theta(s^*)\left[u^s_\theta(b^k_{\xi^{s^*},\pi^*})-\pi^*(s^*,b^k_{\xi^{s^*},\pi^*})\right].$$

Moreover, for each $k\in\mathcal{K}$ it holds:

$$\sum_{\theta\in\Theta}\mu_\theta\Big[\phi_\theta(s_1)\left[u^k_\theta(b^k_{\xi^{s_1},\pi})+\pi(s_1,b^k_{\xi^{s_1},\pi})\right]+\phi_\theta(s_2)\left[u^k_\theta(b^k_{\xi^{s_2},\pi})+\pi(s_2,b^k_{\xi^{s_2},\pi})\right]\Big]=$$
$$\sum_{\theta\in\Theta}\mu_\theta\phi^*_\theta(s^*)\left[u^k_\theta(b^k_{\xi^{s^*},\pi^*})+\pi^*(s^*,b^k_{\xi^{s^*},\pi^*})\right].$$

Hence, noticing that for each signal $s\in\mathcal{S}\setminus\{s_1,s_2\}$ the seller's utility and the buyer's utility does not change from (ϕ,π,p) to (ϕ^*,π^*,p), the set $\mathcal{R}$ of participating buyer's types does not change. As a consequence, the two protocols achieve the same seller's expected utility.

Applying this procedure to all the couples of signals that induces the same vector of best responses, we obtain a generalized-direct and generalized-persuasive protocol providing the same seller's expected utility. □

As a second crucial step, we show that we can focus on protocols without menus (ϕ,p,π) whose price p is equal to the budget b_k of one buyer's type $k\in\mathcal{K}$. Formally:

Lemma J.2. *Given a protocol without menus, there always exists another protocol (ϕ,p,π) such that $p=b_k$ for some $k\in\mathcal{K}$, while achieving the same seller's expected utility as the original protocol.*

Proof. Let (ϕ,π,p) be a protocol. If no type participates, use an uninformative signal, price $p_0=\min_k b_k$, and payment p_0 for every action. The net transfer is zero and all types choose their prior best response, preserving the original outside-option payoff. Hence assume that at least one type participates. We show that there exists a $\hat{k}\in\mathcal{K}$ and a payment function $\hat{\pi}$ such that the protocol $(\phi,\hat{\pi},b_{\hat{k}})$ provides the same seller's expected utility. Let

$$\hat{k}\in\arg\min_{k\in\mathcal{R}_{\phi,\pi,p}:b_k\geq p}\{b_k\}.$$

We observe that all the buyer's types $k\in\mathcal{R}_{\phi,\pi,p}$ have enough budget to participate in the protocol,*i.e.*, $b_k\geq b_{\hat{k}}$. Set $\hat{p}=b_{\hat{k}}$. Furthermore, we define $\hat{\pi}(s,a)=\pi(s,a)+b_{\hat{k}}-p$ for each $s\in\mathcal{S}$ and $a\in\mathcal{A}$.

Then, we show that the set of types $\mathcal{R}_{\phi,\pi,p}=\mathcal{R}_{\phi,\hat{\pi},\hat{p}}$. Indeed, for each type $k\in\mathcal{R}_{\phi,\hat{\pi},\hat{p}}$ it holds

$$\begin{aligned}\sum_{\theta\in\Theta}\sum_{s\in\mathcal{S}}\mu_\theta\phi_\theta(s)\left[u^k_\theta(b^k_{\xi^s,\hat{\pi}})+\hat{\pi}(s,b^k_{\xi^s,\hat{\pi}})\right]-b_{\hat{k}}&\\ =\sum_{\theta\in\Theta}\sum_{s\in\mathcal{S}}\mu_\theta\phi_\theta(s)\left[u^k_\theta(b^k_{\xi^s,\hat{\pi}})+\pi(s,b^k_{\xi^s,\hat{\pi}})+b_{\hat{k}}-p\right]-b_{\hat{k}}&\\ =\sum_{\theta\in\Theta}\sum_{s\in\mathcal{S}}\mu_\theta\phi_\theta(s)\left[u^k_\theta(b^k_{\xi^s,\pi})+\pi(s,b^k_{\xi^s,\pi})\right]-p,&\end{aligned}$$

and hence $k\in\mathcal{R}_{\phi,\pi,p}$. Similarly, we can prove that each buyer's type $k\notin\mathcal{R}_{\phi,\hat{\pi},\hat{p}}$ does not belong to $\mathcal{R}_{\phi,\pi,p}$. It follows that $\mathcal{R}_{\phi,\pi,p}=\mathcal{R}_{\phi,\hat{\pi},\hat{p}}$.

Finally, we can show that the seller's utility results equal to the one in (ϕ,π,p). Indeed, we have:

$$\begin{aligned}&\sum_{k\in\mathcal{R}_{\phi,p,\pi}}\lambda_k\Big[\sum_{\theta\in\Theta}\sum_{s\in\mathcal{S}}\mu_\theta\phi_\theta(s)\left[u^s_\theta(b^k_{\xi^s,\pi})-\pi(s,b^k_{\xi^s,\pi})\right]+p\Big]+\sum_{k\notin\mathcal{R}_{\phi,p,\pi}}\lambda_k\sum_{\theta\in\Theta}\mu_\theta u^s_\theta(b^k_\mu)\\ &\quad=\sum_{k\in\mathcal{R}_{\phi,\hat{p},\hat{\pi}}}\lambda_k\Big[\sum_{\theta\in\Theta}\sum_{s\in\mathcal{S}}\mu_\theta\phi_\theta(s)\left[u^s_\theta(b^k_{\xi^s,\hat{\pi}})-\hat{\pi}(s,b^k_{\xi^s,\hat{\pi}})\right]+b_{\hat{k}}\Big]+\sum_{k\notin\mathcal{R}_{\phi,\hat{p},\hat{\pi}}}\lambda_k\Big[\sum_{\theta\in\Theta}\mu_\theta u^s_\theta(b^k_\mu)\Big]\end{aligned}$$

This concludes the proof. □

Finally, equipped with Lemma J.1 and Lemma J.2, we are ready to provide our polynomial-time algorithm. Intuitively, since we can restrict the attention to protocols without menus (ϕ,p,π) that are generalized-direct and generalized-persuasive, and whose prices p belong to the set $\{b_k\}_{k\in\mathcal{K}}$, we can solve the seller's problem by iterating over all the possible price values $p\in\{b_k\}_{k\in\mathcal{K}}$ and, for each of

them, over all the possible subsets $\mathcal{R} \subseteq \mathcal{K} \cap \{k \in \mathcal{K} : b_k \geq p\}$ of buyer's types that satisfy the IR constraint. This can be done in polynomial time since the number of buyer's types n is fixed. Then, for every $p \in \{b_k\}_{k\in\mathcal{K}}$ and set $\mathcal{R} \subseteq \mathcal{K} \cap \{k \in \mathcal{K} : b_k \geq p\}$, it is sufficient to solve the following problem:

$$\sup_{\substack{\phi\geq 0\\ \pi\geq 0}} \sum_{k\in\mathcal{R}} \lambda_k \left[p + \sum_{\boldsymbol{a}\in\mathcal{A}^n} \sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) \left[u^s_\theta(a_k) - \pi(\boldsymbol{a}, a_k) \right] \right] + \sum_{k\notin\mathcal{R}} \lambda_k \sum_{\theta\in\Theta} \mu_\theta u^s_\theta(b^k_\mu) \quad \text{s.t.} \tag{11a}$$

$$\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a})[u^k_\theta(a_k) + \pi(\boldsymbol{a}, a_k)] \geq \sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a})[u^k_\theta(a') + \pi(\boldsymbol{a}, a')]$$
$$\forall k \in \mathcal{K},\ \forall \boldsymbol{a} \in \mathcal{A}^n,\ \forall a' \in \mathcal{A} \tag{11b}$$

$$\sum_{\boldsymbol{a}\in\mathcal{A}^n} \sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a})[u^k_\theta(a_k) + \pi(\boldsymbol{a}, a_k)] - p \geq \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\mu) \qquad \forall k \in \mathcal{R} \tag{11c}$$

$$\sum_{\boldsymbol{a}\in\mathcal{A}^n} \sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a})[u^k_\theta(a_k) + \pi(\boldsymbol{a}, a_k)] - p \leq \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\mu) \forall k \notin \mathcal{R},\ b_k \geq p \tag{11d}$$

$$\sum_{\boldsymbol{a}\in\mathcal{A}^n} \phi_\theta(\boldsymbol{a}) = 1 \qquad \forall \theta \in \Theta. \tag{11e}$$

By using techniques similar to those used in Section 3 for protocols with menus, we can show that Problem (11) is solvable in polynomial time by means of a suitable-defined LP. Thus, we have:

Theorem J.1. *Restricted to instances in which the number of buyer's types n is fixed, the problem of computing a seller-optimal protocol without menus admits a polynomial-time algorithm.*

Proof. In the following, we present an algorithm to compute an optimal protocol that works in polynomial time when the number of buyer's types is fixed. As a first step, we observe that, thanks to Lemma J.2, the initial payment required by the seller coincides with b_k for some $k \in \mathcal{K}$. Furthermore, we can focus on direct protocols by Lemma J.1. Similarly to Section 3, we can provide a linear relaxation of Problem (11) introducing a variable $l(\boldsymbol{a}, a')$ that replaces $\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) \pi(\boldsymbol{a}, a')$ for each $\boldsymbol{a} \in \mathcal{A}^n$ and $a' \in \mathcal{A}$. Then, we obtain the following LP.

$$\max_{\phi\geq 0, l\geq 0} \sum_{k\in\mathcal{R}} \lambda_k \left[p + \sum_{\boldsymbol{a}\in\mathcal{A}^n} \left[\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) u^s_\theta(a_k) - l(\boldsymbol{a}, a_k) \right] \right] + \sum_{k\notin\mathcal{R}} \lambda_k \sum_{\theta\in\Theta} \mu_\theta u^s_\theta(b^k_\mu) \quad \text{s.t.} \tag{12a}$$

$$\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) u^k_\theta(a_k) + l(\boldsymbol{a}, a_k) \geq \sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) u^k_\theta(a') + l(\boldsymbol{a}, a')$$
$$\forall k \in \mathcal{K},\ \forall \boldsymbol{a} \in \mathcal{A}^n,\ \forall a' \in \mathcal{A} \tag{12b}$$

$$\sum_{\boldsymbol{a}\in\mathcal{A}^n} \left[\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) u^k_\theta(a_k) + l(\boldsymbol{a}, a_k) \right] - p \geq \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\mu) \qquad \forall k \in \mathcal{R} \tag{12c}$$

$$\sum_{\boldsymbol{a}\in\mathcal{A}^n} \left[\sum_{\theta\in\Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) u^k_\theta(a_k) + l(\boldsymbol{a}, a_k) \right] - p \leq \sum_{\theta\in\Theta} \mu_\theta u^k_\theta(b^k_\mu) \forall k \notin \mathcal{R},\ b_k \geq p \tag{12d}$$

$$\sum_{\boldsymbol{a}\in\mathcal{A}^n} \phi_\theta(\boldsymbol{a}) = 1 \qquad \forall \theta \in \Theta. \tag{12e}$$

Write $F(\phi, l)$ for its objective function. Every protocol normalized by Lemmas J.1 and J.2, with price p and participating set $\mathcal{R}$, gives a feasible point of this LP with the same value. Each feasible LP has a bounded objective, since $F(\phi, l) \leq p + 1$, and therefore attains its maximum.

To compute the optimal protocol we can iterate over all the possible prices $p \in \{b_k\}_{k\in\mathcal{K}}$ and all the possible subsets $\mathcal{R} \subseteq \{k \in \mathcal{K} : b_k \geq p\}$ of participating types, solve LP (12), discard infeasible LPs, and return the solution with highest value. Moreover, it is easy to check that the overall procedure requires to solve $O(n2^n)$ LPs, showing that the algorithm runs in polynomial time.

To conclude the proof, we need to show how to modify the solution of LP (12) to obtain a protocol with at least the same value. If $\sum_\theta \mu_\theta \phi_\theta(\bar{\boldsymbol{a}}) = 0$, Constraint (12b) implies that $l(\bar{\boldsymbol{a}}, \bar{a}_k)$ is the maximum payment variable at $\bar{\boldsymbol{a}}$ for every $k \in \mathcal{K}$. Denote this common value by $l(\bar{\boldsymbol{a}})$. Set all payment variables at $\bar{\boldsymbol{a}}$ to zero and add $l(\bar{\boldsymbol{a}})$ to the payment variable of every action at any signal $\hat{\boldsymbol{a}}$ with positive probability. This preserves all best-response inequalities, every type's expected utility, the exclusion inequalities, and the objective. After applying this procedure to all zero-probability signals, set, for every $a' \in \mathcal{A}$,

$$\pi(\boldsymbol{a}, a') = \begin{cases} \dfrac{l(\boldsymbol{a}, a')}{\sum_{\theta \in \Theta} \mu_\theta \phi_\theta(\boldsymbol{a})}, & \text{if } \sum_{\theta \in \Theta} \mu_\theta \phi_\theta(\boldsymbol{a}) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

We observe that the recommended actions are best responses. If the buyer chooses a best response other than the recommended action, their expected utility is unchanged, while seller-favorable tie-breaking guarantees that the seller's expected payoff is at least that from the recommended action. A type can change its participation decision relative to $\mathcal{R}$ only when indifferent; seller-favorable tie-breaking then also weakly increases the seller's payoff. Thus, the recovered protocol has value at least $F(\phi, l)$, even if its participating set differs from $\mathcal{R}$. Every protocol is bounded above by the largest LP value, and recovering a solution attaining that value gives a protocol with at least that utility. Hence the recovered protocol is optimal. This concludes the proof. □

We observe that, thanks to Theorem J.1, the problem of computing a seller-optimal protocol without menus always admits a maximum.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state all the main contributions made by the paper.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: All the assumptions are clearly stated in Section 2.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: All the theoretical results clearly state their assumptions, while all their proofs are provided in the Appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: The paper does *not* include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: The paper does *not* include experiments.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper does *not* include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The paper does *not* include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: The paper does *not* include experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The paper conforms with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: There is no societal impact of the work performed, since the work is mainly theoretical.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper poses no such risks.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: The paper does not use existing assets.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: The paper does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.