

# Persuading a Credible Agent

Jiarui Gan, Abheek Ghosh, Nicholas Teh

University of Oxford, UK

## Abstract

How to optimally persuade an agent who has a private type? When *elicitation* is feasible, this amounts to a fairly standard principal-agent-style mechanism design problem, where the persuader employs a mechanism to first elicit the agent’s type and then plays the corresponding persuasion strategy based on the agent’s report. The optimal mechanism design problem in this setting is relatively well-understood in the literature, with incentive compatible (IC) mechanisms known to be optimal and computationally tractable. In this paper, we study this problem given a *credible* agent, i.e., if the agent claims they are of a certain type in response to the mechanism’s elicitation, then they will act optimally with respect to the claimed type, even if they are actually not of that type.

We present several interesting findings in this new setting that differ significantly from results in the non-credible setting. In terms of the structure of optimal mechanisms, we show that not only may IC mechanisms fail to be optimal, but all mechanisms following the standard ‘eliciting-then-persuading’ mechanism design structure may be suboptimal. To achieve optimality requires two additional instruments—*pre-signaling* and *non-binding elicitation*—which naturally result in multi-stage mechanisms. We characterize optimal mechanisms under these design choices. Based on our characterization, we provide a polynomial-time algorithm for computing optimal multi-stage mechanisms. We also discover that in scenarios that allow for it, *partial* information elicitation can be employed to improve the principal’s payoff even further. Though, surprisingly, an unbounded number of rounds of information exchange between the principal and the agent may be necessary to achieve optimality.

## 1 Introduction

Information design (or *Bayesian persuasion*) models scenarios where a principal strategically reveals her private information to incentivize an agent to play preferred actions. For instance, a seller may want to reveal only partial details about her products to incentivize an agent to buy a particular one, or a school may want to issue only coarse grades to improve its students’ job outcomes [Boleslavsky and Cotton, 2015; Kamenica, 2019; Ostrovsky and Schwarz, 2010].

The standard information design model assumes that the principal knows the utility function of the agent [Kamenica and Gentzkow, 2011], which may not hold in practice. Numerous works subsequently study a model where the principal does not know the agent’s type but has a prior over it [Alonso and Câmara, 2018; Bernasconi et al., 2023; Castiglioni et al., 2020, 2022, 2023; Gill and SgROI, 2008; Kolotilin et al., 2017; Perez-Richet, 2014]. The principal may try to design the signaling scheme (i.e., strategically reveal her private information) to maximize her expected utility given only this prior. Another, possibly better, alternative for the principal, if feasible, is to ask the agent to reveal his type before the principal sends her signal, through an *elicitation* process. In this case, the agent may be *credible*, i.e., take actions consistent with his reported type, or be *non-credible*, i.e., report any type and then play any, possibly inconsistent, action afterward

(essentially, cheap talk). In this paper, we study such scenarios, with a particular focus on credible agents.

Even though credibility has been relatively understudied in the literature, it remains crucial in many situations. One such scenario is when the principal and the agent both enjoy and want to protect some form of reputation. This may happen, for example, when the two parties are known businesses. The principal may still have private information about the state and the agent about his private type, but the agent may stick with the contract (type) signed by both the parties at an earlier stage of the game. Besides reputation, the agent may also want to stick to the previously settled type (contract) to avoid possible legal ramifications. From a behavioral game-theoretic perspective, agents may also be averse to playing actions that are inconsistent with their report due to psychological and social reasons that cause bounded rationality. In strategic scenarios, credibility may also arise voluntarily when an agent intends to behave consistently to change the principal’s belief about their behavior. Such phenomenon have been termed *impersonation* or *imitative deception* in various studies [Gan et al., 2019b; Kash and Parkes, 2010]. Our work is an attempt to capture these situations and understand the computational and algorithmic aspects of mechanism design against credible agents.

We aim to compute payoff-maximizing mechanisms for the principal. In pursuit of this, we identify several crucial design aspects: whether the principal uses only incentive-compatible mechanisms; whether the principal sends other signals in addition to the action recommendations; whether the agent reports other information in addition to his credible type report; and, in general, how many rounds of communication between the principal and the agent are necessary? As it turns out, the presence of credibility changes the design of optimal mechanisms significantly. It necessitates additional design choices that would be unnecessary when the agent is non-credible. We introduce novel stages in the mechanism design and provide non-trivial examples to demonstrate strict payoff improvements they achieve. Moreover, we characterize optimal mechanisms under these additional design choices and prove the general optimality they lead to. The characterization also result in an efficient algorithm to compute optimal mechanisms, thereby highlighting both the computational and utility benefits they offer. Additionally, we provide matching hardness results that emerge when these additional design choices are not employed, thereby highlighting their computational benefits.

## 1.1 Our Results

In Section 3, we begin with a non-credible agent and observe that incentive compatible (IC) policies are without loss of optimality. Given this, we can compute the optimal policy in polynomial time using a linear programming (LP) formulation. For a credible agent, we show that if the principal restricts herself to using only IC policies, then again she can compute the optimal IC policy in polynomial time using an LP. However, optimal IC policies may not be optimal overall, in fact, can be arbitrarily worse; we provide an example that demonstrates this. We then show that computing an optimal policy exactly or approximately is NP-hard using a tight reduction from MAXIMUM INDEPENDENT SET.

In Section 4, we go beyond two-stage interactions (of elicitation and signaling) and consider cases where the principal and the agent may interact over multiple rounds. For a non-credible agent, such interactions beyond two stages are redundant. However, with credible agents, we show that two additional stages—*pre-signaling* and *non-binding elicitation*—are necessary for achieving optimality. We also show via non-trivial examples that these additional stages can strictly improve the principal’s utility. Surprisingly, these additional stages also make the computation of optimal mechanisms tractable. To derive this result, we characterize the class of optimal mechanisms

and prove that a four-stage mechanism structure—involving (in order) pre-signaling, non-binding elicitation, binding type elicitation, and signaling—results in IC across multiple stages and is w.l.o.g. optimal against a credible agent. The characterization can be viewed as a revelation principle, where non-binding elicitation indeed reveals the agent’s true type, despite them eventually imitating another agent type. Based on the simple fixed structure from the characterization, we are able to formulate the computation of optimal mechanisms as an LP, which yields a polynomial-time algorithm.

Finally, in Section 6, we observe that a specific form of elicitation which we call *partial information elicitation* can be used to further improve the principal’s utility. We show that under these information elicitation structures, an unbounded number of rounds of information exchange between the principal and the agent may be necessary to achieve optimality. We prove a tight  $\Theta(n)$  bound on the depth of optimal mechanisms, when nature has  $n$  possible states.

Our results highlight sharp contrasts between the computational and algorithmic aspects of credible versus non-credible cases.

## 1.2 Related Work

We build on the standard Bayesian persuasion model of Kamenica and Gentzkow [2011], which was proposed to study the effects of (the provision of) information on the outcomes of strategic interactions. The model has attracted widespread interest in both economics and theoretical computer science (e.g., [Babichenko and Barman, 2017; Dughmi and Xu, 2016; Feng et al., 2024; Kolotilin et al., 2017; Xu, 2020]; also refer to the surveys by Dughmi [2017] and Kamenica [2019]).

Models with private agent type but without credibility has been well-motivated in the literature, in various forms of principal-agent problems (e.g., [Conitzer and Sandholm, 2002; Myerson, 1982]). This is in general related to the economics concept of cheap talk [Farrell and Rabin, 1996]. Several works look into the Bayesian persuasion model where agents have a private type and reports it (possibly strategically) to the principal. Castiglioni et al. [2022] look into computational problems in a model where the agent is asked to report their type to principal after the principal have committed to a menu of signaling schemes. Subsequent work by Bernasconi et al. [2023] look into further computational questions in this model. Kolotilin et al. [2017] consider a model where the agent only needs to choose between two actions, and both the principal’s and agent’s utilities are linearly related to the state and the agent’s type. They ask the question of whether the principal can benefit by designing a complex mechanism under these assumptions. Other works that consider models of Bayesian persuasion where the agent has a private type include [Gill and SgROI, 2008; Perez-Richet, 2014] and [Alonso and C  mara, 2018]. In *online* Bayesian persuasion [Castiglioni et al., 2020, 2023], the agent has an unknown private type, and this type is chosen adversarially at each round, over multiple principal-agent interactions. Such two-way information exchanges may also take place in sequential interactions as studied more recently in [Gan et al., 2023].

Another related line of work is strategic communication with costly lying [Kartik et al., 2007; Kartik, 2009; Kephart and Conitzer, 2015, 2016; Nguyen and Tan, 2021] or constraints on how they can lie [Yu, 2011; Zhang et al., 2021], which is somewhat the opposite of the *cheap talk* model [Crawford and Sobel, 1982]. However, these works focus mostly on direct cost or constraints imposed on the principal’s signaling or the agent’s reporting. The credibility constraint in our model can be viewed as an implicit cost that depends on both the agent’s report and their action: an infinitely large cost is incurred if they act in ways different from their report.

Our model of a credible agent is related to various studies on *impersonation* or *imitative deception* in algorithmic game theory [Bimpas et al., 2021; Gan et al., 2019a,b; Kash and Parkes, 2010; Nguyen and Xu, 2019]. In these studies, an agent imitates the best responses of another

agent type to change the principal’s belief about their incentives. Such attempts can be effective especially when the principal learns the agent’s incentives by querying the agent’s best responses [Blum et al., 2014; Peng et al., 2019]. From the principal’s perspective, to maximize the utility of her learning outcome amounts to the same mechanism problem we study in this paper. Gan et al. [2019b] considered this mechanism design problem in normal-form Stackelberg games. In their model, the agent (follower) responds to the principal’s (leader’s) strategy according to the utility function they report to the principal, so as to trick the principal into permanently believing that they are of the reported type. Despite the similarity, they did not consider the broader set of approaches we study in this paper. Indeed, a key difference is that in their model, the principal does not have any information advantages over the agent (there is no hidden state of nature, and the principal only determines an action to play). Without such advantages, the standard mechanism structure, which maps the type the agent reports to an outcome, suffices for achieving optimality.

There are other recent works focusing on cheap-talk style conversations in Bayesian settings, where multiple agents with their own private information engage in conversations, through which they strategically reveal their private information [Mao et al., 2022; Leme et al., 2023; Arieli et al., 2023]. Such conversations may proceed in multiple rounds, similarly to multiple stages of information revelation and elicitation in the mechanism we design, but for different fundamental reasons. Crucially, all information exchanges in these models are cheap, in the sense that they do not directly change the players’ payoffs, whereas in ours the agent’s reporting is costly because of credibility: the agent will suffer an infinite cost if he behaves differently from the type he reports.

## 2 Model and Preliminaries

For a set  $S$ , let  $\Delta(S)$  denote the set of all probability distributions over  $S$ . Let  $[n] = \{1, 2, \dots, n\}$  for a positive integer  $n \in \mathbb{Z}_{>0}$ . In the standard information design model [Kamenica and Gentzkow, 2011], a principal (sender) wants to persuade an agent (receiver) to take an action that is desirable to the principal. There is a set of actions  $A$ , and the payoff of each action  $a \in A$  to both the principal and the agent is determined by the state of nature  $\theta \in \Theta$ —we use  $v(\theta, a)$  and  $u(\theta, a)$  to denote those payoffs, respectively. We assume  $\theta$  is drawn from a common prior distribution  $\mu$ , and the principal must commit to a signaling scheme  $\pi(\cdot | \theta) \in \Delta(G)$ , where  $G$  denotes some set of signals. For this standard model, the revelation principle tells that, without loss of generality, the set of signals corresponds to the set of actions, i.e.,  $G = A$ . The goal we adopt from the perspective of a principal is to find an optimal signaling strategy, which maximizes the principal’s expected payoff for the outcome induced.

### 2.1 Agent with Private Type

In many scenarios, the agent has a private utility function unknown to the principal. The standard model can be extended to these scenarios as follows. Let  $T$  be a finite set of types for the agent. An agent of type  $t \in T$  has a utility function  $u_t(\theta, a)$ . The probability that the agent is of type  $t \in T$  is  $\rho(t)$ . The distribution  $\rho \in \Delta(T)$  is known to the principal. The utility function of the principal is  $v(\theta, a)$ , same as the standard model.

When the principal cannot elicit any additional information from the agent about his type, she has to design her signaling scheme  $\pi(\cdot | \theta) \in \Delta(G)$ , for each state  $\theta$ , where  $G$  denotes the set of signals, given only the knowledge about the distribution of types  $\rho$ . Let  $\pi(\theta, g) = \pi(g | \theta) \cdot \mu(\theta)$ . We assume that the agent plays optimally given  $\pi$ . Formally, given  $\pi$ , we assume that an agent of type  $t$  plays an action  $a$  with probability  $\alpha_{\pi, t}(a | g)$  after getting a signal  $g \in G$  from the principal; where

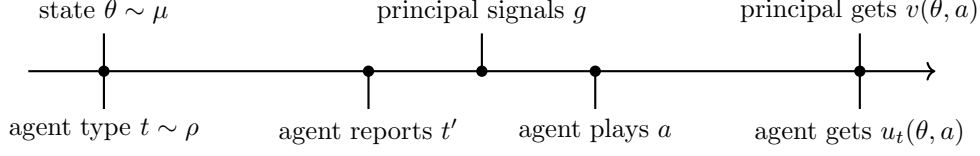


Figure 1: Timeline.

$\alpha_{\pi,t}(a | g) > 0$  only if playing  $a$  maximizes agent's expected utility given  $g$  and  $\pi$ . The expected utility of the agent with type  $t$  is given by

$$\mathbb{E}_{(\theta,g) \sim \pi, a \sim \alpha_{\pi,t}(\cdot | g)}[u_t(\theta, a)] = \sum_{\theta,g} \pi(\theta, g) \sum_a \alpha_{\pi,t}(a | g) \cdot u_t(\theta, a). \quad (1)$$

The expected utility of the principal is given by

$$\mathbb{E}_{t \sim \rho, (\theta,g) \sim \pi, a \sim \alpha_{\pi,t}(\cdot | g)}[v(\theta, a)] = \sum_t \rho(t) \sum_{\theta,g} \pi(\theta, g) \sum_a \alpha_{\pi,t}(a | g) \cdot v(\theta, a). \quad (2)$$

In this model, where the principal cannot elicit any additional information from the agent, [Castiglioni et al. \[2020\]](#) and [Xu \[2020\]](#) prove that the problem of finding a signaling scheme  $\pi$  that maximizes the principal's utility is NP-hard to solve exactly or approximately. In particular, they show that for any constant  $0 \leq \epsilon < 1$ , the principal cannot find a policy that gives her an expected reward of at least  $(1 - \epsilon)$  times the optimal reward in polynomial time unless  $P = NP$ .

## 2.2 Type Elicitation

In this paper, we focus on the scenario with elicitation, i.e., when the principal can ask the agent for his type before sending a signal. The agent of type  $t$  reports his type as  $t' \in T$ . The principal's signaling scheme  $\pi_{t'}(\cdot | \theta)$  is a function of both  $\theta$  and  $t'$ . Let  $\pi_{t'}(\theta, g) = \pi_{t'}(g | \theta) \cdot \mu(\theta)$ . Figure 1 illustrates the timeline of the process.

We consider both *non-credible* and *credible* reporting, focusing primarily on the latter. In the non-credible case, the agent can report any type and then play any action afterward—the action need not be consistent with the reported type. In contrast, in the credible case, the agent needs to behave as if they are the reported type, i.e., if the agent reports a type  $t'$ , then he must play an action that is optimal for the type  $t'$  given his posterior belief (as a function of the common prior and the signal he receives from the principal). We formally introduce the pre-signaling, non-binding elicitation, and partial elicitation in Sections 4 and 6.

## 3 Non-credible vs. Credible Reporting

Let us first compare credible and non-credible reporting and present some algorithmic results.

### 3.1 Non-credible Reporting

When the agent is non-credible, we show that the principal can compute the optimal mechanism in polynomial time using an LP formulation.

First, notice that, when the agent is non-credible, IC is w.l.o.g. using standard arguments: Any mechanism  $\pi$  of the principal will induce an optimal reporting strategy of the agent given  $\pi$ , which

can be encapsulated in a meta-mechanism that simulates the behavior of the agent given his true type. This meta-mechanism plays optimally for the agent and, therefore, is IC. It also provides the same expected reward to the principal as the original mechanism  $\pi$ . So, the principal can optimize over IC signaling policies w.l.o.g. Let  $\pi_t$  denote the mechanism used by the principal when an agent reports type  $t$ . Let  $\pi_t(\theta, a) = \pi_t(a | \theta) \cdot \mu(\theta)$  denote the marginal probability of  $(\theta, a)$ . The following LP computes the optimal signaling mechanism for the principal.

$$\max_{(\pi_t)_{t \in T}} \sum_{t \in T} \rho(t) \sum_{\theta, a} \pi_t(\theta, a) \cdot v(\theta, a) \quad (3)$$

$$\text{subject to } \sum_{\theta, a} \pi_t(\theta, a) \cdot u_t(\theta, a) \geq \sum_{\theta, a} \pi_{t'}(\theta, a) \cdot u_t(\theta, f(a)), \quad \forall t, t' \in T, f : A \rightarrow A, \quad (4)$$

$$\sum_a \pi_t(\theta, a) = \mu(\theta), \quad \forall t \in T, \theta \in \Theta, \quad (5)$$

$$\pi_t \in \Delta(\Theta \times A), \quad \forall t \in T. \quad (6)$$

Specifically, the objective function captures the principal's expected payoff under the agent's truthful response. Constraint (4) ensures that this truthful behavior is indeed incentivized, where: the left side is the agent's expected payoff for truth reporting, conditioned on type  $t$ ; and the right side describes the payoff the agent would obtain if he reports a different type  $t'$  and plays  $f(a)$  when he gets the recommendation to play  $a$ . This constraint also implicitly captures the following:

$$\sum_{\theta} \pi_t(\theta, a) \cdot u_t(\theta, a) \geq \sum_{\theta} \pi_t(\theta, a) \cdot u_t(\theta, a'), \quad \forall t \in T, a, a' \in A,$$

so in addition to truthful type reporting  $\pi$  incentivizes, every action recommendation of  $\pi_t$  is also IC. Since the constraint (4) is defined for all  $f : A \rightarrow A$ , it includes exponentially many linear constraints. However, since  $f$  is not a function of the state  $\theta$ , the constraint can be equivalently written as follows:

$$\sum_{\theta, a} \pi_t(\theta, a) \cdot u_t(\theta, a) \geq \sum_a \underbrace{\max_{a'} \sum_{\theta} \pi_{t'}(\theta, a) \cdot u_t(\theta, a')}_{z(t, t', a)}, \quad \forall t, t' \in T. \quad (7)$$

We can further substitute the maximization on the right side with an auxiliary variable  $z(t, t', a)$  and force the variable to be an upper bound of the maximum value using the following additional constraints, which is linear w.r.t.  $z$  and  $\pi$ :

$$z(t, t', a) \geq \sum_{\theta} \pi_{t'}(\theta, a) \cdot u_t(\theta, a'), \quad \forall t, t' \in T, a, a' \in A. \quad (8)$$

This converts (4) to  $|T|^2$  constraints in (7) and  $|T|^2|A|^2$  constraints in (8), with  $|T|^2|A|$  additional variables  $z(t, t', a)$ . Hence, the size of this new LP formulation is polynomial in the input size of the problem. This gives a polynomial-time algorithm to compute an optimal mechanism against a non-credible agent. This does not appear to follow from the proven properties of these algorithms without further analysis.

### 3.2 Credible Reporting

Now consider a credible agent. After reporting a type  $t'$ , the agent always behaves as if he is actually of type  $t'$ . Formally, say the agent has type  $t$  and reports type  $t'$ . Based on the reported



type  $t'$  and the state  $\theta$ , the principal will send a signal  $g \in G$  with probability  $\pi_{t'}(g | \theta)$ , where  $\pi_{t'}$  is the mechanism that the principal uses for a report  $t'$ ; let  $\pi_{t'}(\theta, g) = \pi_{t'}(g | \theta) \cdot \mu(\theta)$ . After receiving this signal  $g$ , the agent must imitate the reported type  $t'$  to remain credible, so he will select an action  $a'$  that satisfies

$$a' \in \arg \max_{a \in A} \sum_{\theta} \pi_{t'}(\theta, g) \cdot u_{t'}(\theta, a).$$

But notice that the agent receives a payoff based on his actual payoff function  $u_t$  and is equal to  $\sum_{\theta} \pi_{t'}(\theta, g) \cdot u_t(\theta, a')$ .

As with a non-credible agent, the principal can compute the optimal IC mechanism for a credible agent using an LP, although with a slightly different formulation (weaker constraints).

$$\max_{(\pi_t)_{t \in T}} \sum_{t \in T} \rho(t) \sum_{\theta, a} \pi_t(\theta, a) \cdot v(\theta, a) \quad (9)$$

$$\text{subject to } \sum_{\theta, a} \pi_t(\theta, a) \cdot u_t(\theta, a) \geq \sum_{\theta, a} \pi_{t'}(\theta, a) \cdot u_t(\theta, a) \quad \forall t, t' \in T \quad (10)$$

$$\sum_{\theta} \pi_t(\theta, a) \cdot u_t(\theta, a) \geq \sum_{\theta} \pi_t(\theta, a) \cdot u_t(\theta, a') \quad \forall t \in T, a, a' \in A \quad (11)$$

$$\sum_a \pi_t(\theta, a) = \mu(\theta) \quad \forall t \in T, \theta \in \Theta \quad (12)$$

$$\pi_t \in \Delta(\Theta \times A) \quad \forall t \in T \quad (13)$$

Notice that constraint (4) in the LP for a non-credible agent splits into constraints (10) and (11) for a credible agent. Constraint (4) considers all possible deviations over reports and actions of the agent jointly, while (10) and (11) consider such deviations separately. Constraint (10) ensures IC in the reporting stage of the game: it only considers reported-type deviations, because the principal knows that a credible agent is going to play consistently with his reported type. On the other hand, constraint (11) ensures IC in the action stage of the game, by considering only action deviations; this constraint is w.l.o.g. because we can apply the revelation principle to the last stage of the game (when the agent plays his action) and optimize only over policies that are IC (for this stage).

Although the principal can efficiently compute the optimal IC mechanism for credible agents using the LP above, the next example shows that using a non-IC mechanism can sometimes strictly improve her utility. So, the class of IC policies is *not* without loss of optimality for credible agents.

**Example 3.1.** Let there be two agent types  $T = \{t_1, t_2\}$ , two states  $\Theta = \{\theta_1, \theta_2\}$ , and four actions  $A = \{a, b, c, d\}$ . The payoff functions of the agent for the two types are given in the tables below.

	$a$	$b$	$c$	$d$		$a$	$b$	$c$	$d$
$\theta_1$	4	1	3	0	$\theta_1$	0	0	1	0
$\theta_2$	1	4	0	3	$\theta_2$	0	0	0	1
Type $t_1$					Type $t_2$				

The principal gets payoff 1 if the agent plays actions  $c$  or  $d$  and payoff 0 if the agent plays actions  $a$  or  $b$ , irrespective of the state. Let the prior be uniform,  $\mu(\theta_1) = \mu(\theta_2) = 1/2$ .

Let us first consider what happens when a credible agent reports each of the possible types in the above example.

- If the agent reports type  $t_2$ , then by credibility she must act like an agent of type  $t_2$ , irrespective of her actual type. Notice that for type  $t_2$ , the actions in  $\{c, d\}$  dominate those in  $\{a, b\}$ . In particular, for any posterior belief about the state, say  $(p, 1 - p)$ , that the agent may have after receiving principal's signal, he has payoff 0 for playing  $a$  or  $b$  but payoff  $\max(p, 1 - p) > 0$  for playing the better option among  $c$  or  $d$ . So, after reporting type  $t_2$ , the agent always plays an action that gives the principal payoff 1.
- On the other hand, if the agent reports type  $t_1$ , she must always play an action in  $\{a, b\}$  and never an action in  $\{c, d\}$ . In particular, action  $a$  dominates action  $c$ , and action  $b$  dominates  $d$ . As the agent never plays  $c$  or  $d$ , the principal only gets payoff 0.

In an IC mechanism, an agent reports type  $t_i$  only when he is actually  $t_i$ , which happens with probability  $\rho(t_i)$ . Therefore, the expected payoff of the principal is  $0 \times \rho(t_1) + 1 \times \rho(t_2) = \rho(t_2)$ . On the other hand, we propose the following non-IC mechanism that achieves an expected payoff 1, which is strictly greater than  $\rho(t_2)$  if  $\rho(t_1) > 0$ .

The non-IC mechanism works as follows: If the agent reports  $t_1$ , the principal does not give any information about the state to the agent, but if the agent reports  $t_2$ , the principal fully reveals the state. Intuitively, the principal trades information with the agent's "promise" to act as the more preferred type  $t_2$ . This is effective only when the agent is credible. Consider the response of a type- $t_1$  agent under this mechanism.

- If he reports  $t_1$ , as the principal does not reveal any information, the agent's posterior belief is the same as the prior  $(1/2, 1/2)$ , and playing either of  $a$  or  $b$  gives him an expected payoff of  $5/2$ .
- If he reports  $t_2$ , the principal reveals the state, and depending upon whether the state is  $\theta_1$  or  $\theta_2$ , he plays either  $c$  or  $d$  (because he must imitate  $t_2$ ) and gets a payoff of 3.

So, the type- $t_1$  agent is incentivized to report and act like  $t_2$ . It is trivial to check that an agent of type  $t_2$  will report  $t_2$ , too. Consequently, all agent types will report  $t_2$  and play either  $c$  or  $d$ , and the principal gets expected payoff 1. The ratio of between the expected payoffs for the principal for the optimal and optimal-IC mechanism is  $1/\rho(t_2) \rightarrow \infty$  as  $\rho(t_2) \rightarrow 0$ .

Given that IC mechanisms may not be optimal, the principal may want to optimize over all policies, including non-IC mechanisms, too. The theorem below says that even to compute an approximately optimal mechanism is NP-hard. The proof uses a tight reduction from MAXIMUM INDEPENDENT SET, similarly to the approach by Gan et al. [2019b].<sup>1</sup>

**Theorem 1.** *Unless  $P=NP$ , there exists no polynomial-time  $\frac{1}{(|T|-1)^{1-\epsilon}}$ -approximation algorithm for computing a mechanism (against a credible agent) for any constant  $\epsilon > 0$ , even when there are only two possible states of nature.*

In the remainder of the paper, unless otherwise specified all our results assume a credible agent.

## 4 Moving Beyond Two Stages

We have seen that the principal may benefit by using non-IC mechanisms. As it turns out, when credibility is required, there are even more instruments the principal can leverage to further improve her payoff. In this section, we demonstrate two such instruments: (1) pre-signaling, and (2) non-binding elicitation. Pre-signaling means that the principal sends a signal before she elicits the

---

<sup>1</sup>All omitted proofs can be found in the appendix.



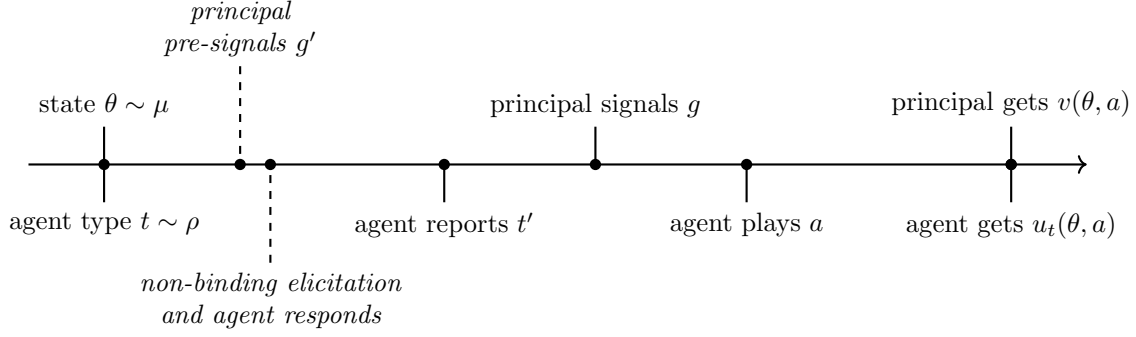


Figure 2: Timeline with pre-signaling and non-binding elicitation.

agent's type. Non-binding elicitation is a cheap-talk communication in addition to the agent's type reporting. It asks the agent how they would like the mechanism to proceed. Since the question is not type related, it is non-binding; no restrictions are attached to the agent's answer on how they must act subsequently.

These additional approaches naturally result in mechanisms involving more stages of interaction between the principal and the agent, beyond the two-stage elicitation-then-signaling paradigm studied in the previous section; see Figure 2. We demonstrate the strict payoff improvement of these mechanisms, their general optimality, and more surprisingly, the computational benefits they bring: in sharp contrast to the intractability of the two-stage mechanisms, we present a polynomial time algorithm to optimize this new type of mechanism.

#### 4.1 Pre-signaling

Intuitively, pre-signaling aims to change the agent's belief to influence their decision on which type to imitate. In the following example, we compare mechanisms with/without pre-signaling, and show that pre-signaling offers a strict improvement in payoff for the principal. For ease of description, we refer to a mechanism that does not use pre-signaling as an ES mechanism, where E and S represents the elicitation and signaling stages, respectively, in the mechanism. Similarly, we refer to a mechanism that uses pre-signaling as an SES mechanism, where the first S refers to the additional pre-signaling stage.

**Example 4.1.** Let  $\Theta = \{\alpha, \beta, \gamma\}$ ,  $T = \{t_0, t_1\}$ , and  $A = \{a, a', b, b'\}$ . The principal's payoff is 1 for actions  $a'$  and  $b'$ , and 0 for actions  $a$  and  $b$ , irrespective of the state. The agent's payoffs are given in the table below. Moreover, the type and state distributions are uniform:  $\rho(t_0) = \rho(t_1) = 1/2$ , and  $\mu(\gamma) = \mu(\alpha) = \mu(\beta) = 1/3$ .

	$a$	$a'$	$b$	$b'$		$a$	$a'$	$b$	$b'$	
$\alpha, \beta, \gamma$	0	1	0	1		$\alpha$	3	2	-2	-3
	$Type\ t_0$					$\beta$	-2	-3	3	2
					$\gamma$	0	-10	0	-10	
						$Type\ t_1$				

persuade a type- $t_1$  agent to imitate  $t_0$ , while there exists an SES mechanism that achieves this with a positive probability.

#### 4.1.1 Without Pre-signaling

When there is no pre-signaling, the key observation (Lemma 2) is that a type- $t_1$  agent will always report truthfully. Since actions  $a'$  and  $b'$  are strictly dominated by  $a$  and  $b$ , a type- $t_1$  agent will then never choose  $a'$  or  $b'$ . As a result, the principal gets payoff 0 when the agent's true type is  $t_1$ , which happens with probability  $\rho(t_1) = 0.5$ . Overall, the principal's expected payoff is at most  $0.5 \times 0 + 0.5 \times 1 = 0.5$  (where 1 is the principal's maximum attainable payoff according to the example).

**Lemma 2.** *In Example 4.1, a type- $t_1$  agent has no incentive to misreport his type as  $t_0$  under any ES mechanism.*

*Proof.* Consider an arbitrary ES mechanism. For a type- $t_1$  agent, even without any additional signal from the principal, they are able to secure an expected payoff of  $1/3$  by sticking to action  $a$ . Now suppose they misreport type  $t_0$ , and we show that this leads to a contradiction.

For a type- $t_0$  agent, actions  $a'$  and  $b'$  strictly dominate  $a$  and  $b$ , respectively. Hence, when an agent acts as  $t_0$ , they will play only  $a'$  or  $b'$ . This means the following for a type- $t_1$  agent for every  $\theta \in \Theta$ :

$$\mathbb{P}(a' | \theta, t_1) + \mathbb{P}(b' | \theta, t_1) = 1,$$

where  $\mathbb{P}$  denotes the probability measure induced by the ES mechanism (and the agent's best response to this mechanism). It follows that the payoff of the type- $t_1$  agent would be:

$$\begin{aligned} \sum_{\theta \in \Theta} \sum_{x' \in \{a', b'\}} \mathbb{P}(x' | \theta, t_1) \cdot u_{t_1}(\theta, x') &= \sum_{\theta \in \Theta} \mu(\theta) \cdot \left( \mathbb{P}(a' | \theta, t_1) \cdot u_{t_1}(\theta, a') + \mathbb{P}(b' | \theta, t_1) \cdot u_{t_1}(\theta, b') \right) \\ &\leq \sum_{\theta \in \Theta} \mu(\theta) \cdot \max \{u_{t_1}(\theta, a'), u_{t_1}(\theta, b')\} = 1/3 \cdot 2 + 1/3 \cdot 2 + 1/3 \cdot (-10) < 0, \end{aligned} \quad (14)$$

which is strictly lower than the minimum payoff  $1/3$  that is guaranteed by reporting truthfully. Therefore, a type- $t_1$  agent is always better off reporting truthfully.  $\square$

#### 4.1.2 With Pre-signaling

Now, consider the following mechanism that uses pre-signaling. Suppose that, before the agent reports their type, the principal pre-signals  $g_\gamma$  when the state is  $\gamma$ , and pre-signals  $g_{\neg\gamma}$  when it is any other state. Depending on the pre-signal sent, the game proceeds as follows:

1. When  $g_\gamma$  is sent, the agent knows with certainty that the state is  $\gamma$ .
2. When  $g_{\neg\gamma}$  is sent, the agent's belief is a uniform distribution over  $\alpha$  and  $\beta$ .

In case 2, by using the following signaling strategy, the principal can obtain payoff 1 subsequently:

$$\pi_{t_1}(a | \alpha) = \pi_{t_1}(a | \beta) = 1, \quad \text{and} \quad \pi_{t_0}(a' | \alpha) = \pi_{t_0}(b' | \beta) = 1.$$

Specifically,  $\pi_{t_1}$  is uninformative as it always recommends  $a$ , whereas  $\pi_{t_0}$  completely differentiates  $\alpha$  and  $\beta$ . Under this mechanism, agents of both types will be incentivized to always (mis)report  $t_0$  and perform the recommended actions. The principal obtains payoff 1 as a result. Overall, since the marginal probability of case 2 is  $\mu(\alpha) + \mu(\beta) = 2/3$ , the principal can obtain payoff at least  $2/3$  in expectation. This is a strict improvement compared with 0.5 in the case without pre-signaling.

The analysis in Sections 4.1.1 and 4.1.2 then implies the following.

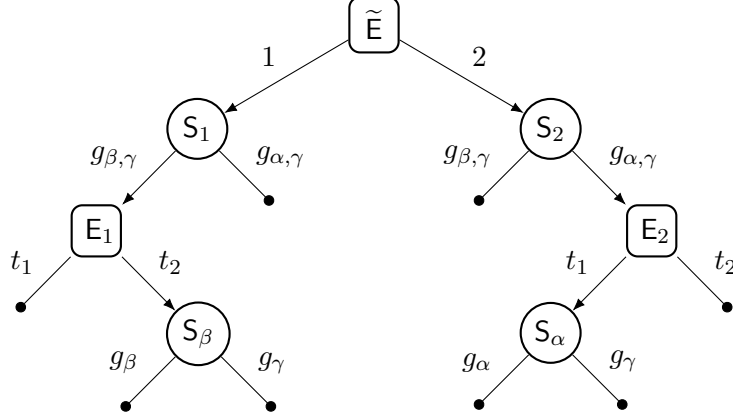


Figure 3: A mechanism that uses non-binding elicitation.

**Theorem 3.** *There exists an instance where an optimal SES mechanism yields a strictly higher payoff for the principal than any ES mechanism does.*

## 4.2 Non-binding Elicitation

Non-binding elicitation can be thought of as a menu for the agent to choose how they would like the mechanism to proceed. An example of mechanisms with non-binding elicitation is illustrated in Figure 3 (we will shortly demonstrate that it offers strict payoff improvement in Example 4.2). In this mechanism, a non-binding elicitation stage is arranged at the beginning of the mechanism. We will use the notation  $\tilde{E}$  to represent non-binding elicitation. At  $\tilde{E}$ , the principal asks the agent to choose between two options labelled 1 and 2, respectively. Depending on the agent's choice, the principal will subsequently proceed with one of the SES mechanisms. The elicitation at  $\tilde{E}$  does not involve any type information and is therefore *non-binding* by nature: the agent's report at  $\tilde{E}$  does not result in any credibility constraints on how they should behave subsequently.

Following our previously established notation, we will refer to such a mechanism as an  $\tilde{E}$ SES mechanism. We demonstrate in the following example that there exists an  $\tilde{E}$ SES mechanism that strictly outperforms any SES mechanism.

**Example 4.2.** *Let  $\Theta = \{\alpha, \beta, \gamma\}$ ,  $T = \{t_1, t_2\}$ , and  $A = \{a, a', b, b', d\}$ . Irrespective of the state, the principal's payoff is 1 for actions  $a'$  and  $b'$ , and 0 for all the other actions. The agent's payoffs are given in the tables below, where all blank entries are  $-M$ , for some sufficiently large constant, say  $M > 10000$ . The type and state distributions are uniform:  $\mu(\alpha) = \mu(\beta) = \mu(\gamma) = 1/3$  and  $\rho(t_1) = \rho(t_2) = 1/2$ .*

	$a$	$a'$	$b$	$b'$	$d$
$\alpha$		1			0
$\beta$			3	2	0
$\gamma$		-1			0

Type  $t_1$

	$a$	$a'$	$b$	$b'$	$d$
$\alpha$	3	2			0
$\beta$				1	0
$\gamma$				-1	0

Type  $t_2$

#### 4.2.1 With Non-binding (and Binding) Elicitation

We first present an  $\tilde{\text{ESES}}$  mechanism that yields a payoff strictly greater than  $1 - \delta$  for the principal in Example 4.2, where  $\delta = 10/M$ . The mechanism is illustrated in Figure 3. The non-binding elicitation at the beginning of the mechanism offers two options 1 and 2. The signaling strategies used at the signaling nodes  $S_1$ ,  $S_3$ ,  $S_\alpha$ , and  $S_\beta$  are defined as follows.

- We use  $\pi_1$  at  $S_1$ , where:

$$\pi_1(g_{\alpha,\gamma} | \alpha) = 1, \quad \pi_1(g_{\beta,\gamma} | \beta) = 1, \quad \text{and} \quad \begin{cases} \pi_1(g_{\alpha,\gamma} | \gamma) = 1 - \delta \\ \pi_1(g_{\beta,\gamma} | \gamma) = \delta \end{cases}.$$

Namely,  $\pi_1$  is designed in a way that induces the following posterior beliefs about the state when the agent receives  $g_{\alpha,\gamma}$  and  $g_{\beta,\gamma}$ :

$$\mathbb{P}(\cdot | g_{\alpha,\gamma}, S_1) = \left( \frac{1}{2-\delta}, 0, \frac{1-\delta}{2-\delta} \right) \quad \text{and} \quad \mathbb{P}(\cdot | g_{\beta,\gamma}, S_1) = \left( 0, \frac{1}{1+\delta}, \frac{\delta}{1+\delta} \right), \quad (15)$$

where  $\mathbb{P}$  denotes the probability measure induced by the mechanism and the agent's best response. The states in the subscript of each signal correspond to states in the support of the posteriors. It will also be useful to note that the marginal probability of each signal is:

$$\mathbb{P}(g_{\alpha,\gamma} | S_1) = (2 - \delta)/3, \quad \text{and} \quad \mathbb{P}(g_{\beta,\gamma} | S_1) = (1 + \delta)/3.$$

- The strategy  $\pi_2$  used at  $S_2$  is defined symmetrically: we swap  $\alpha$  with  $\beta$  in the definition of  $\pi_1$  to obtain  $\pi_2$ .
- The strategies  $\pi_\alpha$  and  $\pi_\beta$  used at  $S_\alpha$  and  $S_\beta$  are truth-revealing strategies:  $\pi_\alpha(g_\theta | \theta) = \pi_\beta(g_\theta | \theta) = 1$  for all  $\theta \in \{\alpha, \beta, \gamma\}$ .

At each leaf node, the mechanism terminates and the agent takes an action that complies with the credibility constraint. If no type elicitation has taken place along the path from the root to the leaf node, the agent is free to take an optimal action of any type—without loss of generality, they will just take an optimal action according to their actual type.

The mechanism ensures the principal a payoff of  $1 - \delta$ , which is close to her payoff upper bound 1 in this example. Indeed, to ensure this high payoff, the agent needs to be persuaded to play  $a'$  or  $b'$  almost all the time. Examining each state separately, we will find that the incentives of types  $t_1$  and  $t_2$  align with the this goal at states  $\alpha$  and  $\beta$ , respectively, but not at the other states. Hence, we need to incentivize a type- $t_1$  agent to act as  $t_2$  with a sufficiently high probability at state  $\beta$ , and incentivize a type- $t_2$  agent to act as  $t_1$  at state  $\alpha$ . This is what the mechanism does: it keeps  $\alpha$  and  $\beta$  mixed with  $\gamma$  (at  $S_1$  and  $S_2$ , respectively); subsequently, it “disentangles”  $\alpha$  and  $\beta$  from  $\gamma$  (at  $S_\alpha$  and  $S_\beta$ , respectively) only if the agent promises to act as  $t_2$  and  $t_1$ , respectively.

The non-binding elicitation plays a role here because the ways we mix and disentangle the states need to be designed differently for the two agent types. Indeed, the following lemma shows that the design incentivizes each type  $t_i$  to choose the corresponding option  $i$  at  $\tilde{\text{E}}$ .

**Lemma 4.** *In Example 4.2, given the mechanism presented in Figure 3, for each  $i \in \{1, 2\}$  the optimal strategy of a type- $t_i$  agent is to select option  $i$  at  $\tilde{\text{E}}$ .*

Following the lemma, by symmetry, the principal obtains the same payoff on  $t_1$  and  $t_2$ . Consider the situation against a type- $t_1$  agent. It can be verified that (see the proof of Lemma 4 for more details):

- When  $g_{\alpha,\gamma}$  is sent, the agent will react optimally by playing action  $a'$ , and the principal obtains payoff 1.
- When  $g_{\beta,\gamma}$  is sent, the agent acts as  $t_2$ , playing  $b'$  when it is  $\beta$  and  $d$  when it is  $\gamma$ . The principal obtains payoff  $\mathbb{P}(\beta | g_{\beta,\gamma}, S_1) \cdot 1 = 1/(1 + \delta)$ .

Overall, the principal's payoff is

$$\mathbb{P}(g_{\alpha,\gamma} | S_1) \times 1 + \mathbb{P}(g_{\beta,\gamma} | S_2) \times \frac{1}{1 + \delta} = \frac{2 - \delta}{3} \times 1 + \frac{1 + \delta}{3} \times \frac{1}{1 + \delta} > 1 - \delta.$$

#### 4.2.2 With Binding Elicitation Only

Consider an arbitrary SES mechanism  $\Pi$ ,<sup>2</sup> which only uses binding elicitation. We will argue that  $\Pi$  yields payoff at most  $1 - \delta$  for the principal. In what follows, let  $\pi : \Theta \rightarrow \Delta(G)$  be the signaling strategy used in the pre-signaling step of  $\Pi$ , such that  $\pi(g) > 0$  for all  $g \in G$ . Let  $\mathbb{P}$  denote the probability measure induced by  $\Pi$  (and the agent's optimal response) and let  $\mathbb{E}$  denote the expectation over  $\mathbb{P}$ .

The proof of the  $1 - \delta$  upper bound relies on the following lemmas. Lemma 5 first argues that if  $\Pi$  were to achieve a high payoff for the principal, there must be a high-payoff signal  $g$  that implies  $\gamma$  with a sufficiently high probability. (Recall that  $v$  denotes the principal's payoff function.) Lemma 6 further shows that the principal must receive a high-payoff signal that implies both  $\alpha$  and  $\beta$  with a sufficiently high probability. With these observations, we prove the  $1 - \delta$  upper bound on the performance of all SES mechanisms in Example 4.2 and establish Theorem 7.

**Lemma 5.** *Suppose that  $\mathbb{E}(v) \geq 1 - \delta$  in Example 4.2. Then, for any  $\epsilon \in (0, 1 - \delta)$ , there exists a signal  $g \in G$  such that  $\mathbb{E}(v | g) > 1 - \delta - \epsilon$  and  $\mathbb{P}(\gamma | g) \geq 1/4 - \delta/\epsilon$ .*

*Proof sketch.* The converse of the statement asserts the existence of  $\epsilon$  such that: for every  $g \in G$ , if  $\mathbb{P}(\gamma | g) \geq 1/4 - \delta/\epsilon$  then  $\mathbb{E}(v | g) \leq 1 - \delta - \epsilon$ . Indeed, since the prior probability is  $\mu(\gamma) = 1/4$ , the marginal probability of signals  $g$  such that  $\mathbb{P}(\gamma | g) \geq 1/4 - \delta/\epsilon$  cannot be insignificant. Hence, if the converse statement is true, then for a substantial portion of signals would yield expected payoff lower than  $1 - \delta - \epsilon$  for the principal, which would further lead to an overall payoff strictly lower than the assumed value  $1 - \delta$ .  $\square$

**Lemma 6.** *Suppose that  $\mathbb{E}(v) \geq 1 - \delta$  in Example 4.2. Then there exists a signal  $g \in G$  such that  $\mathbb{E}(v | g) > 1 - 11\delta$ ,  $\mathbb{P}(\alpha | g) > 1/100$ , and  $\mathbb{P}(\beta | g) > 1/100$ .*

*Proof sketch.* By Lemma 5 some high-payoff  $g$  must imply  $\gamma$  with a significant probability. Now that there is no non-binding elicitation, both agent types will receive  $g$  with some probability. For type  $t_1$  (and similarly for type  $t_2$  and state  $\beta$ ), the probability of  $\alpha$  in the posterior conditioned on  $g$  must be sufficiently high: otherwise either  $b$  or  $d$  will always strictly dominate other actions (e.g., consider the extreme case where  $\mathbb{P}(\alpha | g) = 0$ ), leading to a low payoff 0 for the principal and a contradiction to the assumption that  $\mathbb{E}(v | g) > 1 - 11\delta$ .  $\square$

**Theorem 7.** *There exists an instance where an optimal  $\tilde{\text{ESES}}$  mechanism yields a strictly higher payoff for the principal than any SES mechanism.*

<sup>2</sup>We will hereafter mostly use  $\Pi$  to denote a multi-stage mechanism and use  $\pi$  to denote the signaling strategy in  $\Pi$  for the pre-signaling stage.

*Proof.* We show that in Example 4.2 no SES mechanism yields payoff higher than  $1 - \delta$  for the principal. Consider an arbitrary SES mechanism and suppose for the sake of contradiction that it yields payoff  $\mathbb{E}(v) > 1 - \delta$  for the principal.

Note that the mechanism elicits the agent's type in the second step. Consider the case where the agent is of type  $t_1$  and a signal  $g$  that satisfies the conditions in Lemma 6 is sent. Subsequently, when the mechanism elicits the agent's type, the agent either reports  $t_1$  or  $t_2$ . Consider each of these two possible reports. We show that both cases lead to contradictions to complete the proof.

**Case 1:  $t_1$  is reported.** In this case, the type- $t_1$  agent will act truthfully. Consider the agent's expected payoff  $\mathbb{E}[u_{t_1} | t_1, g]$  and the following upper bound on the payoff:

$$\mathbb{E}[u_{t_1} | t_1, g] \leq \mathbb{P}(\beta, a \vee a' | t_1, g) \cdot (-M) + (1 - \mathbb{P}(\beta, a \vee a' | t_1, g)) \cdot 3.$$

It must be  $\mathbb{E}[u_{t_1} | t_1, g] \geq 0$  because the agent can stick to action  $d$  to secure payoff 0. This gives  $\mathbb{P}(\beta, a \vee a' | t_1, g) \leq \frac{3}{M+3}$ . Note that, for type  $t_1$ , action  $b'$  is strictly dominated by  $b$  or  $d$  under any state distribution, which means  $\mathbb{P}(\beta, b' | t_1, g) = 0$ . Therefore,

$$\mathbb{P}(\beta, b \vee d | t_1, g) = \mathbb{P}(\beta | t_1, g) - \mathbb{P}(\beta, a \vee a' | t_1, g) > \frac{1}{100} - \frac{3}{M+3}.$$

Since the principal gets payoff 0 when  $b$  or  $d$  is played, we get that

$$\begin{aligned} \mathbb{E}(v | g) &\leq 1 - \mathbb{P}(b \vee d | g) \leq 1 - \mathbb{P}(t_1, \beta, b \vee d | g) \\ &= 1 - \mathbb{P}(t_1 | g) \cdot \mathbb{P}(\beta, b \vee d | t_1, g) \leq 1 - \frac{1}{3} \cdot \left( \frac{1}{100} - \frac{3}{M+3} \right) < 1 - 11\delta \end{aligned}$$

(note that  $\mathbb{P}(t_1 | g) = \mathbb{P}(t_1) = \rho(t_1) = 1/3$  by independence). This contradicts Lemma 6.

**Case 2:  $t_2$  is reported.** In this case, the agent will act according to type  $t_2$ . We can replicate the above arguments to show that  $\mathbb{P}(\alpha, a \vee d | t_1, g) > \frac{1}{100} - \frac{3}{M+3}$ , where we replace  $\beta$  with  $\alpha$ ,  $a$  and  $a'$  with  $b$  and  $b'$ , respectively, and  $u_{t_1}$  with  $u_{t_2}$ . Eventually, this also leads to the contradiction  $\mathbb{E}(v | g) < 1 - 11\delta$ .  $\square$

## 5 Computing Optimal Multi-stage Mechanisms

$\tilde{\text{ESES}}$  and  $\text{SES}$  mechanisms are strictly better than  $\text{ES}$  mechanisms in terms of the payoffs they generate. But what do they imply computationally? Is it easier or harder to compute optimal  $\text{ESES}$  and  $\text{SES}$  mechanisms? As it turns out, optimal  $\text{SES}$  mechanisms are still inapproximable by noting that in the reduction for Theorem 1 the principal does not benefit from using pre-signaling (Theorem 8). But optimal  $\tilde{\text{ESES}}$  mechanisms, while generating even higher payoffs than  $\text{SES}$ s, are efficiently computable. We next present an efficient algorithm for  $\tilde{\text{ESES}}$  and demonstrate its general optimality among a broad class of multi-stage mechanisms.

**Theorem 8.** *Unless  $P = NP$ , there exists no polynomial-time  $\frac{1}{(|T|-1)^{1-\epsilon}}$ -approximation algorithm for computing an  $\text{SES}$  mechanism for any constant  $\epsilon > 0$ , even when there are only two possible states of nature.*

*Proof.* We use the same reduction as that in the proof of Theorem 1 and note that the principal does not benefit from using pre-signaling in the reduced instance.  $\square$



## 5.1 The General Optimality of $\tilde{\text{ESES}}$

The efficient algorithm relies on a characterization of optimal  $\tilde{\text{ESES}}$  mechanisms. The characterization is quite general: it compares  $\tilde{\text{ESES}}$  mechanisms with all possible multi-stage mechanisms involving any number of signaling and binding or non-binding elicitation steps. We hereafter refer to this broader class of mechanisms as *indefinite-stage mechanisms*. Similarly to the mechanisms we presented previously, an indefinite-stage mechanism can be history-dependent, specifying different strategies for different nodes of the game tree. For binding elicitation, we consider only direct elicitation, where the principal asks the agent directly “what is your type?” and the agent must choose one type in  $T$  as their answer to this question. (In the next section, we will further investigate a more general type of elicitation that allows the principal to elicit partial type information.)

For ease of description, in what follows, we assume that the type set contains  $n$  types:  $T = \{t_1, \dots, t_n\}$ . Since an  $\tilde{\text{ESES}}$  mechanism corresponds to a tree as we illustrated in Figure 3, we will describe the mechanism based on the tree. The tree contains the following types of nodes:

- **S nodes.** An S node represents a signaling step, where the principal sends a signal according to a strategy  $\pi : \Theta \rightarrow \Delta(G)$  for this node. The node then has  $|G|$  child nodes each corresponding to a distinct signal in  $G$ . When a signal  $g$  is drawn, the mechanism proceeds to the child corresponding to  $g$ .
- **E nodes.** An E node represents to a binding elicitation step, where the agent reports their type. Each E node has  $n$  children, each corresponding to a type in  $T$ . When the agent reports a type  $t$ , the mechanism proceeds to the child corresponding to  $t$ .
- **$\tilde{\text{E}}$  nodes.** An  $\tilde{\text{E}}$  node represents to a non-binding elicitation step, where the agent selects a child of this node, and the mechanism proceeds to the selected child. The root node of an  $\tilde{\text{ESES}}$  mechanism is always an  $\tilde{\text{E}}$  node.
- **Leaf nodes.** At a leaf node, the mechanism terminates, and the agent performs an optimal action of the *reported type*. If no type elicitation takes place on the path to the leaf, we assume that the agent is free to choose an optimal action of any type.<sup>3</sup>

To sidestep intricate corner cases, we follow the convention and assume that if there is a tie in the agent’s decision-making (whether to decide an action to play or a type to report), the mechanism decides how to break the tie.

The characterization result, presented in the theorem below can be viewed as a generalized revelation principle that involves multiple levels of IC or DIC in an  $\tilde{\text{ESES}}$  mechanism (where “D” emphasizes *direct* recommendation). Intuitively, the option the agent selects at  $\tilde{\text{E}}$  reveals their true type, though this does not mean that the agent needs to subsequently act truthfully because of the non-binding nature of this elicitation. Subsequently, the signaling strategy at each  $S_i$  directly recommends the agent a type to report, while those at  $S_{ij}$  directly recommend actions.

**Theorem 9.** *For any indefinite-stage mechanism  $\Pi$ , there exists an  $\tilde{\text{ESES}}$  mechanism  $\Pi'$  that yields as much payoff for the principal as  $\Pi$  does and has the following properties.*

- IC at  $\tilde{\text{E}}$ :** *The root  $\tilde{\text{E}}$  has  $n$  children  $S_1, \dots, S_n$ , which are all S nodes. Each type- $t_i$  agent is incentivized to choose the branch leading to  $S_i$ .*

---

<sup>3</sup>W.l.o.g., the agent will choose an optimal action of their actual type. In fact, it is w.l.o.g. to assume that there is always one E node on every path from the root to a leaf (see the proof of Theorem 9).

- ii. **DIC at  $S_i$ :** Each node  $S_i$  has  $n$  E node children  $E_{i1}, \dots, E_{in}$ . Each  $E_{ij}$  is associated with type  $t_j$ , and at  $E_{ij}$  a type- $t_i$  agent is incentivized to report  $t_j$ . For each  $k = 1, \dots, n$ , reporting type  $t_k$  at  $E_{ij}$  leads to an S node child  $S_{ijk}$ .
- iii. **DIC at  $S_{ijk}$ :** The signaling strategy used at each node  $S_{ijk}$  directly recommends the agent an action to take, and the action is optimal for type  $t_k$ .

*Proof sketch.* Given any  $\Pi$ , we construct an  $\tilde{\text{ESES}}$  mechanism  $\Pi'$  that ensures the stated properties while preserving the principal's payoff by following the steps below.

- *Step 1. Adding a new root.* We make  $n$  copies  $\Pi^1, \dots, \Pi^n$  of  $\Pi$  and attach each  $\Pi^i$  as a subtree to the root of  $\Pi'$ . Let the root node be an  $\tilde{E}$  node. Since all the copies are the same, each type  $t_i$  is incentivized to choose  $\Pi^i$  and we get IC at the root.
- *Step 2. Removing  $\tilde{E}$  nodes.* Given IC at the root, in each branch  $\Pi^i$ , we only need to consider the response of  $t_i$ . Hence, all the  $\tilde{E}$  nodes in  $\Pi^i$  become redundant: we remove each  $\tilde{E}$  node  $e$  and reconnect the child of  $e$  selected by a type- $t_i$  agent to the parent of  $e$ .
- *Step 3. Merging S nodes.* Next, we merge consecutive S nodes on every path as if signals generated in these consecutive signaling processes are generated at once: i.e., we treat the combination of the signals as a meta-signal sent by the joint process.

The above operations yield a tree with five levels containing only  $\tilde{E}$ , S, E, S, and leaf nodes, respectively. We proceed as follows to achieve DIC at the S nodes.

- *Step 4. Ensuring DIC.* The idea is to merge signals incentivizing the same response of the agent (i.e., reporting the same type or taking the same action) into a single signal. The signal can be viewed as a recommendation to the agent (in terms of what to report and which action to take). Essentially, these merging operations will make many nodes in  $\Pi$  indistinguishable to the agent, but if the agent follows the recommendations in the new mechanism, they would still reach each leaf with the same probability as before, as if they respond optimally in  $\Pi$ . On the other hand, since these merging operations reduce information, the agent cannot devise any better response to improve their payoff. Hence, the agent is incentivized to follow the recommendations and this ensures DIC.  $\square$

## 5.2 An LP Formulation

The characterization in Theorem 9 fixes the structure of mechanism. The only parameters to be optimized are the strategies at the S nodes, including  $\pi_i : \Theta \rightarrow \Delta(T)$  at each  $S_i$ , and  $\pi_{ijk} : \Theta \rightarrow \Delta(A)$  at each  $S_{ijk}$ , where the signal spaces of the strategies are determined by the DIC properties. At a high level, this results in an *extensive-form game*. We formulate the problem of optimizing the principal's commitment in this game as an LP. The variables of the LP are as follows; each of them corresponds to the marginal probabilities of a path on the game tree, and the feasible space of the variables can be characterized by a set of flow constraints.

- $\pi_i(t_j | \theta)$  for  $i, j \in \{1, \dots, n\}$ , which represents the strategy  $\pi_i$ .
- $\phi_{ijk}(a | \theta)$  for  $i, j \in \{1, \dots, n\}$  and  $a \in A$ , which represents the product  $\pi_i(g_j | \theta) \cdot \pi_{ijk}(g_a | \theta)$  and indirectly represents  $\pi_{ijk}(g_a | \theta)$ . We use these variables instead of  $\pi_{ijk}(g_a | \theta)$  to avoid quadratic terms in our formulation.

The objective of the LP is to maximize the expected payoff of the principal yielded by the mechanism. Under the IC and DIC conditions, this can be expressed as follows (where we highlight the variables in blue):

$$\max \sum_{\theta \in \Theta} \sum_{i=1}^n \sum_{j=1}^n \sum_{a \in A} \rho(t_i) \cdot \mu(\theta) \cdot \phi_{ijj}(a | \theta) \cdot v(a, \theta),$$

Namely, the expression assume that each type  $t_i$  is incentivized to select option  $i$  at  $\tilde{E}$ , report  $t_j$  at each  $E_{ij}$ , and play the action recommended by each  $\pi_{ijj}$ . These conditions are enforced next through the constraints of the LP:

- The flow constraint ensures the product encoded in  $\phi_{ijk}$  correspond to a feasible  $\pi_{ijk}$ :

$$\begin{aligned} \sum_{a \in A} \phi_{ijk}(a | \theta) &= \pi_i(t_j | \theta) && \text{for all } i, j, k \in \{1, \dots, n\} \text{ and } \theta \in \Theta \\ \phi_{ijk}(a | \theta) &\geq 0 && \text{for all } i, j, k \in \{1, \dots, n\}, a \in A, \text{ and } \theta \in \Theta \end{aligned}$$

We also have that  $\pi_i(\cdot | \theta) \in \Delta(T)$  for all  $i$  and  $\theta$ , which ensure that  $\pi_i$  is a valid strategy.

- The following constraint ensures DIC at each  $S_{ijk}$ , i.e., it is optimal for the *imitated type*  $t_k$  to play the recommended action  $a$ , rather than any other action  $b$ :

$$\begin{aligned} \sum_{\theta} \mu(\theta) \cdot \phi_{ijk}(a | \theta) \cdot u_{t_k}(a, \theta) &\geq \sum_{\theta} \mu(\theta) \cdot \phi_{ijk}(b | \theta) \cdot u_{t_k}(b, \theta) \\ &\text{for all } a, b \in A \text{ and } i, j, k \in \{1, \dots, n\}. \end{aligned}$$

Namely, the left side is the marginal payoff of a type- $t_k$  agent for playing action  $a$  when  $a$  is recommended, and the right side is that for playing any other action  $b$ .

- Given the above constraint, we introduce the following additional variable  $u_{t_i}(S_{\ell jk})$  to capture a type- $t_i$  agent's maximum attainable payoff at each  $S_{\ell jk}$ :

$$u_{t_i}(S_{\ell jk}) = \sum_{a \in A} \sum_{\theta} \mu(\theta) \cdot \phi_{\ell jk}(a | \theta) \cdot u_{t_i}(a, \theta).$$

The following constraint then ensures DIC at each  $E_{ij}$ , i.e., it is optimal for an agent whose *actual type* is  $t_i$  to report the recommended type  $t_j$ , rather than any other type  $t_k$ :

$$u_{t_i}(S_{ijj}) \geq u_{t_i}(S_{ijk}) \quad \text{for all } i, j, k \in \{1, \dots, n\}.$$

- Finally, the following constraint ensures IC at  $\tilde{E}$ , so that each type  $t_i$  is incentivized to select option  $i$ , rather than any other option  $\ell$ :

$$\sum_{j=1}^n u_{t_i}(S_{ijj}) \geq \sum_{j=1}^n \max_{k=1, \dots, n} u_{t_i}(S_{\ell jk}) \quad \text{for all } i, \ell \in \{1, \dots, n\}. \quad (16)$$

On the right side, the payoff is yield when the agent selects option  $\ell$  and subsequently reports an optimal type  $t_k$  at each node  $E_{\ell j}$ ; hence, there is a maximization operator over  $k$ . As in Section 3, the constraint can be linearized by further introducing an auxiliary variable  $\bar{u}_{t_i}(E_{\ell j})$ , along with constraints  $\bar{u}_{t_i}(E_{\ell j}) \geq u_{t_i}(S_{\ell jk})$  for all  $k$ , to capture an upper bound on  $\max_{k=1, \dots, n} u_{t_i}(S_{\ell jk})$ . With this the right side of (16) can be replaced with  $\bar{u}_{t_i}(S_{\ell jk})$ .

Given the LP, the tractability of optimal indefinite-stage mechanisms then follows readily.

**Theorem 10.** *An optimal indefinite-stage mechanism can be computed in polynomial time.*

## 6 Further Improvement: Partial Information Elicitation

In the mechanisms we have looked at so far, a binding elicitation directly queries the agent’s type. In some scenarios, elicitation might be implemented in a less direct format. For example, the principal may ask the agent “are you type  $t_1$  or not?”, or present several subsets of types and ask the agent which subset contains the agent’s type. After a subset is reported, the principal may follow up with more questions to identify smaller subsets containing the agent’s type (that they want to imitate) until eventually pinning down the type in a singleton. Essentially, this splits one canonical elicitation step into multiple steps interleaved with singling steps. We refer to such elicitation as *partial information elicitation* (PIE) and mechanisms using partial elicitation as *PIE mechanisms*.

PIE is still *binding*, so that once an agent identifies a subset containing his true type, he needs to subsequently report and behave in consistency with this subset. This is crucial; indeed, such partial elicitation does not provide any added value when the agent is non-credible, as we demonstrated previously. In the next example, we demonstrate that PIE can strictly improve the principal’s payoff (when the agent is credible).

### 6.1 Strict Payoff Improvement by Partial Elicitation

**Example 6.1.** Let  $\Theta = \{\alpha, \beta\}$ ,  $T = \{t_0, t_1, t_2\}$ , and  $A = \{a, a', b, b'\}$ . The principal gets payoff 1 for the state-action pairs  $(\alpha, a')$  and  $(\beta, b')$ , and 0 for all other pairs. The agent’s payoffs are given in the table below, where all blank entries are  $-1$ . The state is distributed uniformly:  $\mu(\alpha) = \mu(\beta) = 1/2$ . The type distribution is:  $\rho(t_0) = 1$  and  $\rho(t_1) = \rho(t_2) = 0$  (i.e., only  $t_0$  actually appears).<sup>4</sup>

	$a$	$a'$	$b$	$b'$
$\alpha$	1	0	-2	-3
$\beta$	-2	-3	1	0
Type $t_0$				

	$a$	$a'$	$b$	$b'$
$\alpha$		1		
$\beta$		1		
Type $t_1$				

	$a$	$a'$	$b$	$b'$
$\alpha$				1
$\beta$				1
Type $t_2$				

#### 6.1.1 With PIE Mechanism

Consider the following PIE mechanism. The principal first asks the agent whether their type is  $t_0$ .

- If the answer is “yes”, then no information will be sent subsequently, and the agent decides an action to take. (By credibility, this action needs to be optimal for  $t_0$ .)
- If the answer is “no”, then the principal will send a signal that reveals the true state, and then ask the agent to identify their type from the set  $T \setminus \{t_0\}$  (given that the agent has denied that their type is  $t_0$ ). Finally, the agent performs an optimal action of the type they selected.

It can be verified that the above mechanism yields payoff 1 for the principal, which is the maximum possible payoff of the principal in this example. Specifically (note that we only need to consider a type- $t_0$  agent in this case because the other types have probability 0):

- A type- $t_0$  agent will deny that their type is  $t_0$  in response to the principal’s first question: If they reported  $t_0$ , no information would be provided, in which case states  $\alpha$  and  $\beta$  would still be equally likely according to their belief. Consequently, the agent’s expected payoff would be negative for all possible actions.

<sup>4</sup>For simplicity, we use zero-probability types in this example. One can also construct a similar but more sophisticated example where all types have non-zero probabilities.

- Subsequently, when the principal reveals  $\alpha$ , the agent will be incentivized to report and act as  $t_1$ , responding with  $t_1$ 's optimal action  $a'$ . When the principal reveals  $\beta$ , the agent will act as  $t_2$  and respond with  $t_2$ 's optimal action  $b'$ . In both states, the agent gets payoff 0 (strictly higher than the above case), and the principal gets payoff 1.

### 6.1.2 With Non-PIE Mechanism

Now consider an arbitrary non-PIE (indefinite-stage) mechanism  $\Pi$ . We show that  $\Pi$  yields payoff lower than 1 for the principal.

**Lemma 11.** *Any non-PIE mechanism  $\Pi$  yields payoff lower than 1 for the principal in Example 6.1.*

*Proof.* Suppose for the sake of contradiction that  $\Pi$  yields payoff 1 for the principal. W.l.o.g.  $\Pi$  is an  $\tilde{\text{ESES}}$  mechanism as we argued in Section 5.

Since in the example the probability  $t_1$  and  $t_2$  actually appear is zero, only a type- $t_0$  agent would contribute to the principal's payoff. Consider the distribution induced by  $\Pi$  and a type- $t_0$  agent's optimal response over tuples  $(\theta, t) \in \Theta \times T$ , where  $t$  is the type the agent reports. Let  $(\theta, t)$  be an arbitrary tuple with a positive probability in this distribution. It must be that  $t \neq t_0$ . Indeed, if the agent acts according to  $t_0$ , they will play  $a$  or  $b$  as the other actions are strictly dominated. Since the principal's payoff is 0 for both  $a$  and  $b$ , the principal will get 0 with a positive probability, contradicting the assumption that  $\Pi$  yields payoff 1 for the principal.

As a result, we have  $t \in \{t_1, t_2\}$ . If  $t = t_1$ , it must be  $\theta = \alpha$ . Indeed, a type- $t_1$  agent can only be incentivized to play  $a'$  according to the payoff definition, whereas the principal's payoff for  $(\beta, a')$  is 0. Consider further the binding elicitation step (which is not partial elicitation as  $\Pi$  is non-PIE) and the agent's posterior belief  $p \in \Delta(\Theta)$  at this step. There are two possible cases:

- $p(\alpha) = 1$ . In this case, a type- $t_0$  agent would in fact be better off reporting  $t_0$  and playing  $a$ , instead of reporting  $t_1$  (and playing  $a'$ ); the payoff is 1 for the former and 0 for the latter. Hence,  $p(\alpha) = 1$  is not possible.
- $p(\alpha) < 1$  (and hence  $p(\beta) > 0$ ). In this case, since  $p(\beta) > 0$ , with a positive probability the state-action pair  $(\beta, a')$  will be realized (given the assumption that the agent will act as  $t = t_1$  subsequently). This means that the principal will receive payoff 0 with a positive probability, which contradicts the assumption that  $\Pi$  yields payoff 1 for the principal, too.

By symmetry, the case where  $t = t_2$  also leads to contradictions. The assumption that  $\Pi$  yields payoff 1 for the principal does not hold.  $\square$

The above result and the PIE mechanism we presented above establish the following result.

**Theorem 12.** *There exists an instance where an optimal PIE mechanism yields a strictly higher payoff for the principal than any non-PIE mechanisms do.*

## 6.2 Suboptimality of Constant-Step PIE Mechanisms

It might be tempting to think that, similarly to the case of non-PIE mechanisms, it is without loss of optimality to consider PIE mechanisms with a constant depth. Surprisingly, as we demonstrate using the following example, this is not the case: in this example, the depths of *all* the optimal PIE mechanisms grow with the size of the instance.

	$a_1$	$a_2$	$\dots$	$a_{n-1}$	$a_n$	$a'_1$	$a'_2$	$\dots$	$a'_{n-1}$	$a'_n$	
$\theta_1$	1					0					
$\theta_2$		1			0		0		$-M$		
$\vdots$			$\ddots$					$\ddots$			
$\theta_{n-1}$				0	1		$-M$		0		
$\theta_n$						1				1	

$Type\ t_0$

	$a_1$	$\dots$	$a_i$	$\dots$	$a'_i$	$\dots$
$\theta_1$	0					
$\vdots$						
$\theta_i$				0	$1/M$	
$\vdots$						

$Type\ t_i$

Table 1: The agent’s payoffs in Example 6.2. The payoffs of an agent of type  $t_i$  (for each  $i = 1, \dots, n$ ) is the same as those of type  $t_0$ , with the exception of the  $i + 1$  entries highlighted in gray (payoffs in the blank entries are all the same as type  $t_0$ ).

**Example 6.2.** Let  $\Theta = \{\theta_1, \dots, \theta_n\}$ ,  $T = \{t_0, t_1, \dots, t_n\}$ , and  $A = \{a_1, \dots, a_n\} \cup \{a'_1, \dots, a'_n\}$ . The principal gets payoff 1 for the state-action pairs  $(\theta_i, a'_i)$  for all  $i = 1, \dots, n$ , and gets 0 for all other pairs. The agent’s payoffs are given in Table 1, where we let  $M = 2n^{3^n}$ . The parameters are non-negative only on the diagonal from  $(\theta_1, a_1)$  to  $(\theta_n, a_n)$  and in the entry  $(\theta_n, a'_n)$ . The state is distributed uniformly:  $\mu(\theta_1) = \dots = \mu(\theta_n) = 1/n$ . The type distribution is:  $\rho(t_0) = 1$  and  $\rho(t_1) = \dots = \rho(t_n) = 0$  (i.e., only  $t_0$  actually appears).

### 6.2.1 An Optimal PIE Mechanism

An  $n$ -step PIE mechanism is illustrated in Figure 4. The mechanism begins with an E node,  $E_0$ , and asks the agent whether they are of type  $t_i$  or not at each subsequent node  $E_i$ : answering “yes” leads to termination of the mechanism, while answering “no” leads to continuation of the mechanism to node  $S_{i+1}$ . At each  $S_i$ , the signaling strategy signals  $g_i$  deterministically if the state is  $\theta_i$ , and signals  $g_{>i}$ , otherwise.

Recall that  $t_0$  is the only type that appears with a positive probability. It can be verified that a type- $t_0$  agent is (weakly)<sup>5</sup> incentivized to respond as follows:

- Act as  $t_i$  and play action  $a'_i$  upon receiving  $g_i$ , whereby he receives payoff 0. Indeed, according to the construction, we have  $\mathbb{P}(\theta_i | g_i) = 1$ , so the type- $t_0$  agent could only be better off playing  $a_i$ . However, playing  $a_i$  is blocked by the credibility constraint because upon reaching  $S_i$  it must be that the agent has denied that they are of types  $t_0, \dots, t_{i-1}$ ; playing  $a_i$  cannot be justified as an optimal action of any of the remaining types in this case.
- Report  $\neg t_i$  at each  $E_i$ . This can be verified by noting that the agent’s posterior belief at  $E_i$  is a uniform distribution over  $\theta_{i+1}, \dots, \theta_n$ . Any of  $a_{i+1}, \dots, a_n$  is optimal for the type- $t_0$  agent and gives payoff  $1/(n - i)$ . In comparison, if the agent keeps reporting  $\neg t_j$  at the each subsequent  $E_j$  and acts as type  $t_j$  upon receiving  $g_j$  as described above, he will be persuaded to play precisely  $a'_j$  at each  $\theta_j$ , whereby he gets the same overall payoff  $1/(n - i)$  (thanks in particular to the payoff 1 for  $(\theta_n, a'_n)$ ). Hence, it is optimal to report  $\neg t_i$  at each  $E_i$ .

Since action  $a'_i$  is played at every  $\theta_i$ , the principal always gets payoff 1, which is also the maximum possible payoff in this example. Hence, the PIE mechanism is optimal.

<sup>5</sup>One can change the 0’s on the diagonal from  $(\theta_1, a'_1)$  to  $(\theta_{n-1}, a_{n-1})$  to a sufficiently small number to strongly incentivize the agent and, with extra bookkeeping, prove that the example still works.





Figure 4: A optimal mechanism for Example 6.2,  $i \in \{0, 1, \dots, n-1\}$ .  $E_n$  is a terminal node.

Intuitively, the PIE mechanism works by revealing a state  $\theta_i$  only after the agent denies his type being any of  $t_1, \dots, t_{i-1}$ . This is crucial as otherwise the agent would have been better off acting as any arbitrary type in this range, whose optimal action is  $a_i$ . Moreover, eliciting the exact type immediately at any  $E_i$  would not work as that would only encourage the agent to report  $t_i$  and play one of the actions  $a_{i+1}, \dots, a_n$ , leading to payoff 0 for the principal. We next formally analyze the lower bound on the depths of optimal PIE mechanisms.

### 6.2.2 Depth of Optimal PIE Mechanisms

We demonstrate an  $\Omega(n)$  lower bound on the depths of the optimal PIE mechanisms, for a class of instances with  $n$  possible states. Consider an arbitrary optimal PIE mechanism  $\Pi$ . Given the PIE mechanism we presented above, for  $\Pi$  to be optimal, it must yield payoff 1 for the principal. We will show that to achieve this payoff,  $\Pi$  must have  $\Omega(n)$  levels.

Since type  $t_0$  is the only type that actually appears, an argument similar to the proof of Theorem 9 implies that there is no need for  $\Pi$  to use any  $\tilde{E}$  node. We can then assume w.l.o.g. that  $\Pi$  starts with an  $E$  node, and subsequently the levels alternate between  $S$  and  $E$  nodes (with the exception of leaf nodes, which can appear at any level). For each node  $x$  in  $\Pi$ , we define the following sets:

- $T_x \subseteq T$ , which contains the types that are *not* excluded by the agent's answers at the  $E$  nodes on the path from the root to  $x$ .
- $\Theta_x \subseteq \Theta$ , which is the support of the agent's posterior at  $x$ , i.e.,  $\theta \in \Theta_x$  if and only if  $\mathbb{P}(\theta | x) > 0$ , where  $\mathbb{P}$  is the probability measure induced by  $\Pi$ .

To simplify the notation, we write  $p_x$  as the distribution such that  $p_x(\theta) := \mathbb{P}(\theta | x)$  for each  $\theta$ .

The next two lemmas follow by noting that to yield payoff 1 for the principal requires each  $a'_i$  to be played exactly at state  $\theta_i$ , and this behavior can be induced only when the agent acts as type  $t_i$ .

**Lemma 13.**  $|\Theta_x| = 1$  for every leaf node  $x$  in  $\Pi$ .

**Lemma 14.** For every node  $x$  in  $\Pi$ , if  $\theta_i \in \Theta_x$  then  $t_i \in T_x$ .

Moreover, we can prove the following key result to establish a connection between posteriors at each node and their child nodes.

**Lemma 15.** For every  $i = 1, \dots, n-1$  and  $\epsilon \in \mathbb{R}$ , let  $\mathcal{P}_i^\epsilon \subseteq \Delta(\Theta)$  consist of points  $p$  such that:

- i.  $\sum_{j=i}^n p(\theta_j) = 1$ ;
- ii.  $\sum_{j=i+1}^n p(\theta_j) \geq 1 - \frac{1}{n-i+1} - \epsilon$ ; and

iii.  $|p(\theta_j) - p(\theta_n)| \leq \epsilon/2$  for all  $j = i + 1, \dots, n$ .

Let  $\epsilon_k = n^{-3^{n-k}}$  for  $k \in \mathbb{Z}$ . For every non-leaf node  $x$  in  $\Pi$ , if  $t_i \in T_x$  and  $p_x \in \mathcal{P}_i^{\epsilon_k}$ , then there exists  $c \in \text{children}(x)$  such that  $p_c \in \mathcal{P}_i^{\epsilon_{k+1}}$ .

Intuitively, when  $\epsilon$  is small,  $\mathcal{P}_i^\epsilon$  is roughly a hyperplane in  $\Delta(\Theta)$ , where  $p(\theta_{i+1}) = \dots = p(\theta_n)$ . The above lemma states that if type  $t_i$  is not yet excluded at an S node, then at the next level of the tree there must be a node  $c$  at which the posterior stays in  $\mathcal{P}_i^{\epsilon_{k+1}}$ , which is roughly the same as  $\mathcal{P}_i^{\epsilon_k}$ . When  $\epsilon_{k+1}$  is sufficiently small, this further means that the posterior  $p_c$  remains supported on  $\{\theta_{i+1}, \dots, \theta_n\}$  because the probabilities of these states cannot be too close to 0 due to the first condition in the lemma. By Lemma 13, the mechanism cannot terminate at  $c$  in this case, so  $\Pi$  needs to include one more level. Applying this argument repeatedly we find that  $\Pi$  must have  $\Omega(n)$  levels.

**Theorem 16.** *There exists a class of instances where every optimal PIE mechanism has depth  $\Omega(n)$ .*

*Proof.* Consider the instance given by Example 6.2 and the following induction process.

- Since  $\mu$  is uniform in the example, at the root node  $r$  we have  $p_r = \mu \in \mathcal{P}_1^{\epsilon_1}$  and  $t_1 \in T_r = T$ .
- Suppose that at some non-leaf node  $x$ , it holds that  $p_x \in \mathcal{P}_i^{\epsilon_k}$  and  $t_i \in T_x$  for  $i, k < n - 1$ . By Lemma 15,  $p_c \in \mathcal{P}_i^{\epsilon_{k+1}}$  for some  $c \in \text{children}(x)$ . Consider the following two cases.
  - If  $p_c(\theta_i) > 0$ , which means  $\theta_i \in \Theta_c$ , then by Lemma 14 it must be  $t_i \in T_c$ . Hence, there exists a child  $c$  such that  $p_c \in \mathcal{P}_i^{\epsilon_{k+1}}$  and  $t_i \in T_c$ .
  - If  $p_c(\theta_i) = 0$ , then since  $p_c \in \mathcal{P}_i^{\epsilon_{k+1}}$ , by definition we have  $\sum_{j=i+1}^n p_c(\theta_j) = 1$ . In this case, it must be that

$$0 < p_c(\theta_{i+1}) \leq \frac{1}{n-i} + \frac{\epsilon_{k+1}}{2}. \quad (17)$$

Indeed, if  $p_c(\theta_{i+1})$  was larger than the upper bound, we would have  $p_c(\theta_j) > 1/(n-i)$  for all  $j = i + 1, \dots, n$  due to condition (iii) in the definition of  $\mathcal{P}_i^{\epsilon_{k+1}}$ ; this leads to the contradiction  $\sum_{j=i+1}^n p_c(\theta_j) > 1$ . Similarly,  $p_c(\theta_{i+1}) = 0$  would imply  $p_c(\theta_j) < \epsilon_{k+1}$  by condition (iii), and in turn the contradiction  $\sum_{j=i+1}^n p_c(\theta_j) < n \cdot \epsilon_{k+1} < 1$ .

The upper bound implies

$$\sum_{j=i+2}^n p_c(\theta_j) = 1 - p_c(\theta_{i+1}) \geq 1 - \frac{1}{n-i} - \epsilon_{k+1}.$$

Since  $p_c \in \mathcal{P}_i^{\epsilon_{k+1}}$  already implies that  $|p_c(\theta_j) - p_c(\theta_n)| \leq \epsilon_{k+1}/2$  for all  $j = i + 2, \dots, n$ . This means  $p_c \in \mathcal{P}_{i+1}^{\epsilon_{k+1}}$ .

Additionally, the lower bound in (17) implies  $\theta_{i+1} \in \Theta_c$ , and in turn  $t_{i+1} \in T_c$  by Lemma 14.

Therefore, in both cases, we find a child  $c$  such that  $t_\ell \in T_c$  and  $p_c \in \mathcal{P}_\ell^{\epsilon_{k+1}}$  for some  $\ell \leq i + 1$ . Since  $i < n - 1$ , we also have  $|\Theta_c| > 1$  by noting that  $p_c(\theta_j) > 0$  for all  $j = i + 2, \dots, n$  for the same reason that  $p_c(\theta_{i+1}) > 0$  as we argued in (17). Hence,  $c$  is not a leaf node by Lemma 13.

By induction, there exists a non-leaf node at the  $(n - 2)$ -th level. Hence, the depth of  $\Pi$  is  $\Omega(n)$ .  $\square$

The lower bound is tight by a similar argument to Theorem 9 and noting that it is w.l.o.g. to design partial elicitation in a way such that  $T_c \subsetneq T_x$  for any E node  $x$  and  $c \in \text{children}(x)$ . We state the upper bound below and omit the proof.

**Theorem 17.** *For any PIE mechanism  $\Pi$ , there exists a PIE mechanism  $\Pi'$  with depth  $O(n)$  that yields as much payoff for the principal as  $\Pi$  does.*

We remark that the size of the tree of an optimal mechanism may still grow exponentially in its depth. Because of this, we cannot use LP formulations similar to the one in Section 5 to obtain efficient algorithms for computing optimal PIE mechanisms. We leave open this problem. The problem is intriguing in that even when a single partial elicitation node  $x$  is considered, there are exponentially many possible ways to set up the question for the agent as every subset of  $T_x$  may be involved.

## 7 Conclusion

We studied Bayesian persuasion with credible agents. For two-stage—elicitation then signaling—mechanisms, we showed that IC policies are easy to compute but may be suboptimal, while the optimal policy may be NP-hard to compute. We demonstrate that two additional instruments—pre-signaling and non-binding elicitation—lead to mechanisms involving additional stages, can strictly improve the principal’s utility, are necessary for achieving optimality, and can make the problem computationally tractable. Finally, we also show that a specific form of elicitation that we call partial information elicitation can further improve the principal’s utility, but an unbounded number of rounds of information exchange between the principal and the agent may be necessary to achieve optimality. Future directions include analyzing the computational complexity of the optimal PIE mechanisms, which we leave open in the current paper. It would also be interesting to consider credible agents in similar problems beyond persuasion.

## References

- Ricardo Alonso and Odilon Câmara. On the value of persuasion by experts. *Journal of Economic Theory*, 174:103–123, 2018.
- Itai Arieli, Ivan Geffner, and Moshe Tennenholtz. Mediated cheap talk design. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 5456–5463, 2023.
- Yakov Babichenko and Siddharth Barman. Algorithmic aspects of private bayesian persuasion. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik, 2017.
- Martino Bernasconi, Matteo Castiglioni, Andrea Celli, Alberto Marchesi, Francesco Trovò, and Nicola Gatti. Optimal rates and efficient algorithms for online bayesian persuasion. 2023.
- Georgios Birmpas, Jiarui Gan, Alexandros Hollender, Francisco J Marmolejo-Cossío, Ninad Rajgopal, and Alexandros A Voudouris. Optimally deceiving a learning leader in stackelberg games. *Journal of Artificial Intelligence Research*, 72:507–531, 2021.
- Avrim Blum, Nika Haghtalab, and Ariel D. Procaccia. Learning optimal commitment to overcome insecurity. In *Proceedings of the 28th Conference on Neural Information Processing Systems (NIPS)*, pages 1826–1834, 2014.

- Raphael Boleslavsky and Christopher Cotton. Grading standards and education quality. *American Economic Journal: Microeconomics*, 7(2):248–279, 2015.
- Matteo Castiglioni, Andrea Celli, Alberto Marchesi, and Nicola Gatti. Online bayesian persuasion. *Advances in Neural Information Processing Systems*, 33:16188–16198, 2020.
- Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Bayesian persuasion meets mechanism design: Going beyond intractability with type reporting. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2022.
- Matteo Castiglioni, Andrea Celli, Alberto Marchesi, and Nicola Gatti. Regret minimization in online bayesian persuasion: Handling adversarial receiver’s types under full and partial feedback models. *Artificial Intelligence*, 314:103821, 2023.
- Vincent Conitzer and Tuomas Sandholm. Complexity of mechanism design. In *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*, pages 103–110, 2002.
- Vincent Crawford and Joel Sobel. Strategic information transmission. *Econometrica*, 50(6):1431–1451, 1982.
- Shaddin Dughmi. Algorithmic information structure design: A survey. *SIGecom Exchanges*, 15(2):2–24, 2017.
- Shaddin Dughmi and Haifeng Xu. Algorithmic bayesian persuasion. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 412–425, 2016.
- Joseph Farrell and Matthew Rabin. Cheap talk. *Journal of Economic perspectives*, 10(3):103–118, 1996.
- Yiding Feng, Chien-Ju Ho, and Wei Tang. Rationality-robust information design: Bayesian persuasion under quantal response. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 501–546. SIAM, 2024.
- Jiarui Gan, Qingyu Guo, Long Tran-Thanh, Bo An, and Michael Wooldridge. Manipulating a learning defender and ways to counteract. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 8274–8283, 2019a.
- Jiarui Gan, Haifeng Xu, Qingyu Guo, Long Tran-Thanh, Zinovi Rabinovich, and Michael Wooldridge. Imitative follower deception in Stackelberg games. In *Proceedings of the 2019 ACM Conference on Economics and Computation (EC)*, page 639–657, 2019b.
- Jiarui Gan, Rupak Majumdar, Debmalaya Mandal, and Goran Radanovic. Sequential principal-agent problems with communication: Efficient computation and learning. *arXiv preprint arXiv:2306.03832*, 2023.
- David Gill and Daniel Sgroi. Sequential decisions with tests. *Games and Economic Behavior*, 63(2):663–678, 2008.
- Emir Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11(1):249–272, 2019.
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

- Navin Kartik. Strategic communication with lying costs. *The Review of Economic Studies*, 76(4): 1359–1395, 2009.
- Navin Kartik, Marco Ottaviani, and Francesco Squintani. Credulity, lies, and costly talk. *Journal of Economic Theory*, 134(1):93–116, 2007.
- Ian A Kash and David C Parkes. Impersonation strategies in auctions. In *Internet and Network Economics: 6th International Workshop, WINE 2010, Stanford, CA, USA, December 13-17, 2010. Proceedings 6*, pages 492–495. Springer, 2010.
- Andrew Kephart and Vincent Conitzer. Complexity of mechanism design with signaling costs. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 357–365, 2015.
- Andrew Kephart and Vincent Conitzer. The revelation principle for mechanism design with reporting costs. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 85–102, 2016.
- Anton Kolotilin, Tymofiy Mylovanov, Andriy Zapechelnuk, and Ming Li. Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964, 2017.
- Renato Paes Leme, Jon Schneider, and Shuran Zheng. Bayesian conversations. *arXiv preprint arXiv:2307.08827*, 2023.
- Jieming Mao, Renato Paes Leme, and Kangning Wang. Interactive Communication in Bilateral Trade. In *Proceedings of the 13th Innovations in Theoretical Computer Science Conference (ITCS)*, 2022.
- Roger B Myerson. Optimal coordination mechanisms in generalized principal–agent problems. *Journal of mathematical economics*, 10(1):67–81, 1982.
- Anh Nguyen and Teck Yong Tan. Bayesian persuasion with costly messages. *Journal of Economic Theory*, 193:105212, 2021.
- Thanh H. Nguyen and Haifeng Xu. Imitative attacker deception in Stackelberg security games. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 528–534, 2019.
- Michael Ostrovsky and Michael Schwarz. Information disclosure and unraveling in matching markets. *American Economic Journal: Microeconomics*, 2(2):34–63, 2010.
- Binghui Peng, Weiran Shen, Pingzhong Tang, and Song Zuo. Learning optimal strategies to commit to. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*, pages 2149–2156, 2019.
- Eduardo Perez-Richet. Interim bayesian persuasion: First steps. *American Economic Review*, 104(5):469–474, 2014.
- Haifeng Xu. On the tractability of public persuasion with no externalities. In *Proceedings of the Thirty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2708–2727, 2020.
- Lan Yu. Mechanism design with partial verification and revelation principle. *Autonomous Agents and Multi-Agent Systems*, 22:217–223, 2011.

Hanrui Zhang, Yu Cheng, and Vincent Conitzer. Automated mechanism design for classification with partial verification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 5789–5796, 2021.

David Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 681–690, 2006.



## A Omitted Proofs

### A.1 Proof of Theorem 1

**Theorem 1.** *Unless  $P=NP$ , there exists no polynomial-time  $\frac{1}{(|T|-1)^{1-\epsilon}}$ -approximation algorithm for computing a mechanism (against a credible agent) for any constant  $\epsilon > 0$ , even when there are only two possible states of nature.*

*Proof.* We give a tight reduction from MAXIMUM INDEPENDENT SET. Let  $G = (V, E)$  be an undirected graph, an arbitrary instance of MAXIMUM INDEPENDENT SET. Let  $n = |V|$  denote the number of vertices. Let  $N(v) = \{v' \mid (v, v') \in E\}$  denote the set of neighbors of  $v$ . It is known that approximating the MAXIMUM INDEPENDENT SET problem better than a factor of  $1/n^{1-\epsilon}$  for any  $\epsilon > 0$  is NP-hard [Zuckerman, 2006], i.e., there is no polynomial-time algorithm, unless  $P=NP$ , which, given an arbitrary undirected graph  $G$  with a maximum independent set of size of  $k$ , can always tell whether  $G$  has a maximum independent set of size larger or smaller than  $k/n^{1-\epsilon}$ .

The reduction goes as follows: Let  $T = \{t_v : v \in V\} \cup \{t_*\}$ , i.e., there is a type  $t_v$  corresponding to each vertex  $v \in V$  and a special type  $t_*$ . Note that  $|T| = n + 1$ . In the distribution over types  $\rho$ , each  $t_v$  has an equal probability of occurring,  $\rho(t_v) = 1/n$ , but the type  $t_*$  has a zero probability,  $\rho(t_*) = 0$ .<sup>6</sup> Let there be two states of nature,  $\Theta = \{\theta_1, \theta_2\}$ , which can occur with equal probability,  $\mu(\theta_1) = \mu(\theta_2) = 1/2$ . The action set of the agent is  $A = \{a_v : v \in V\} \cup \{b_v : v \in V\} \cup \{a_0, b_0\}$ . The utility function of type  $t_v$  is given in the table below, where  $v' \in N(v)$  is any arbitrary neighbour of  $v$ , and  $v'' \notin N(v)$  is any arbitrary vertex that is not adjacent to  $v$ .

	$a_0$	$b_0$	$a_v$	$b_v$	$a_{v'}, b_{v'}$	$a_{v''}, b_{v''}$
$\theta_1$	$1 + 1/n$	$-1 + 1/n$	1	-1	-n	0
$\theta_2$	$-1 + 1/n$	$1 + 1/n$	-1	1	-n	0

The utility of type  $t_*$  is  $-1$  for actions  $a_0$  and  $b_0$ , irrespective of the state, and 1 for all other actions and states. Irrespective of the state, the payoff of the principal is 0 for  $a_0$  and  $b_0$ , and 1 for all other actions.

For each type  $t_v$ , the better action in  $\{a_0, b_0\}$  strictly dominates all other actions. In particular,  $a_0$  dominates  $a_v$ ,  $a_{v'}$ , and  $b_{v'}$ ;  $b_0$  dominates  $b_v$  (and  $a_{v'}$  and  $b_{v'}$ ); for any posterior, say  $(p, 1 - p)$ , of the agent after getting the signal from the principal, the better action in  $\{a_0, b_0\}$  gets a utility of  $\max(2p - 1 + 1/n, 1 - 2p + 1/n) \geq 1/n > 0$  and dominates  $a_{v''}$  and  $b_{v''}$ . So, if the agent reports his type as  $t_v$ , for any  $v \in V$ , the principal gets a reward of 0.

On the other hand, type  $t_*$  dislikes playing actions  $a_0$  and  $b_0$ , which give him a reward of  $-1$ , but is indifferent between playing any of the other actions, all of which give him a reward of 1. So, the principal gets a reward of 1 if any agent reports his type as  $t_*$ .

Essentially, the principal's overall utility depends upon the number of types  $t_v$  that can be persuaded to report and behave like  $t_*$ . We prove next that the principal's utility is proportional to the size of the maximum independent set in  $G$ . Let us first prove that the principal's optimal utility is at least the size of the maximum independent set divided by  $n$ . Let  $V' \subseteq V$  be any maximum independent set of  $G$ . We show that the following signaling policy yields a utility of  $|V'|/n$  for the principal.

- The agent reports a type  $t_v$  for some  $v \in V$ . In this case, reveal no information.

---

<sup>6</sup>If this is unsatisfactory, it is easy to check that the proof works for  $\rho(t_*) = \delta > 0$  as well, for small  $\delta$ , with more bookkeeping.

- The agent reports  $t_*$ . If the state is  $\theta_1$ , pick a  $v' \in V'$  uniformly at random and recommend action  $a_{v'}$ , i.e., each action  $a_{v'}$  for  $v' \in V'$  is recommended with probability  $1/|V'|$  and other actions are recommended with probability 0. On the other hand, if the state is  $\theta_2$ , recommend each action  $b_{v'}$  for  $v' \in V'$  with probability  $1/|V'|$  and other actions with probability 0.

Given this signaling policy, let us compute the expected reward that an agent of type  $t_v$  gets by reporting specific types in  $T$ :

- By reporting  $t_{v'}$  for any  $v' \in V$ , the agent gets no information from the principal. His posterior is the same as the prior  $(1/2, 1/2)$ , and he gets an expected utility of  $1/n$  by playing either  $a_0$  or  $b_0$ .
- By reporting  $t_*$ , the agent gets recommendation  $a_{v'}$  or  $b_{v'}$  depending upon the state, for some  $v' \in V'$ . If the vertex  $v$  corresponding to the agent's type  $t_v$  is in  $V'$ , then none of the vertices in  $V'$  are neighbors of  $v$ , so the agent gets an expected reward of 1. On the other hand, if  $v \notin V'$ , then there must be at least one vertex in  $V'$  who is a neighbor of  $v$ , so the agent has an expected utility of  $\leq -n/|V'| + (|V'| - 1)/|V'| < 0$ .

The analysis above tells us that a type- $t_v$  agent prefers to report  $t_*$  if and only if  $v \in V'$ . So, the principal gets an expected utility of  $|V'|/n$ .

Conversely, suppose that the size of the maximum independent sets is  $k$ . We show that the maximum payoff the principal can achieve is at most  $k/n$ . Suppose, for the sake of contradiction, the principal obtains  $k'/n$  by using some signaling policy  $\pi$ , where  $k' > k$ . Since the principal gets 0 for each type  $t_v$  unless they misreport  $t_*$ , there must be at least  $k + 1$  of these types who have been persuaded to misreport  $t_*$ . Given that the maximum independent set has size  $k$ , there must be two adjacent vertices  $v_1, v_2 \in V$ , whose corresponding types  $t_{v_1}$  and  $t_{v_2}$  would both misreport  $t_*$ . We show that this lead to a contradiction.

Note that by reporting  $t_v$ , for any  $v \in V$ , and playing  $a_0$  or  $b_0$ , any type can get a reward of  $1/n$ . Hence, to incentivize type- $t_{v_1}$  to misreport  $t_*$ , the state-action pairs  $(\theta_1, a_{v_1})$  and  $(\theta_2, b_{v_1})$  must be selected with a total probability of at least  $1/n$  because these pairs give type- $t_{v_1}$  a utility of 1 while all other pairs give a utility of  $\leq 0$ . Since  $v_1$  and  $v_2$  are neighbors, type- $t_{v_2}$  gets a negative payoff of  $-n$  when these pairs are selected, which results in an overall payoff of at most  $-n \cdot \frac{1}{n} + (1 + \frac{1}{n}) \cdot (1 - \frac{1}{n}) < 0 < \frac{1}{n}$  (where  $1 + \frac{1}{n}$  is the highest attainable payoff for the agent among any action-state pairs). This contradicts the assumption that  $t_{v_2}$  is incentivized to misreport  $t_*$ .

We showed that the optimal expected utility of the principal and the size of the maximum independent set are proportional (equal, except for multiplicative factor of  $1/n$ ). So, if we had a polynomial-time algorithm for approximating principal's optimal expected utility, we could use it to approximate the size of the maximum independent set, which is not possible unless  $P=NP$ .  $\square$

## A.2 Proof of Lemma 4

**Lemma 4.** *In Example 4.2, given the mechanism presented in Figure 3, for each  $i \in \{1, 2\}$  the optimal strategy of a type- $t_i$  agent is to select option  $i$  at  $\tilde{E}$ .*

*Proof.* Consider a type- $t_1$  agent and compare the agent's payoffs for choosing options 1 and 2.

If the type- $t_1$  agent selects option 1:

- Upon receiving  $g_{\alpha, \gamma}$ , the agent's posterior belief is  $\mathbb{P}(\cdot | g_{\alpha, \gamma}, S_1) = \left(\frac{1}{2-\delta}, 0, \frac{1-\delta}{2-\delta}\right)$  (almost evenly distributed over  $\alpha$  and  $\gamma$ ). It is then optimal for the agent to take action  $a'$ , whereby she gets a strictly positive payoff while all other actions gives at most 0.

- Upon receiving  $g_{\beta,\gamma}$ , the agent's posterior belief is  $\mathbb{P}(\cdot | g_{\beta,\gamma}, S_1) = \left(0, \frac{1}{1+\delta}, \frac{\delta}{1+\delta}\right)$ . It is optimal for the agent to report (and imitate)  $t_2$  in response to the elicitation at  $E_1$ , whereby the agent would play  $b'$  upon receiving  $g_\beta$  (when the state can only be  $\beta$ ) and  $d$  upon receiving  $g_\gamma$  (when the state can only be  $\gamma$ ), as if she is of type  $t_2$ . Overall, the type- $t_1$  agent's expected payoff conditioned on  $E_1$  is  $\frac{2}{1+\delta} > 0$ . In comparison, if she reports  $t_1$  and receives no further information to update her belief, the best action would be  $d$ , which only gives payoff 0.

Overall the agent's expected payoff for selecting option 1 is *at least*

$$\underbrace{\frac{2-\delta}{3}}_{\mathbb{P}(g_{\alpha,\gamma} | S_1)} \times 0 \quad + \quad \underbrace{\frac{1+\delta}{3}}_{\mathbb{P}(g_{\beta,\gamma} | S_1)} \times \frac{2}{1+\delta} \quad = \quad \frac{2+2\delta}{3(1+\delta)}.$$

If the type- $t_1$  agent selects option 2:

- Upon receiving  $g_{\alpha,\gamma}$ , the agent's posterior belief is  $\mathbb{P}(g_{\alpha,\gamma} | S_2) = \left(\frac{1}{1+\delta}, 0, \frac{\delta}{1+\delta}\right)$ . Then, truthful reporting is obviously optimal at  $E_2$ , in particular because more information will be provided if the agent reports  $t_1$ . The agent further plays action  $a'$  upon receiving  $g_\alpha$  (when the state can only be  $\alpha$ ), and  $d$  upon receiving  $g_\gamma$  (when the state can only be  $\gamma$ ). The overall payoff conditioned on  $E_2$  is  $\frac{1}{1+\delta}$ .
- Upon receiving  $g_{\beta,\gamma}$ , the agent's posterior belief is  $\mathbb{P}(g_{\beta,\gamma} | S_2) = \left(0, \frac{1}{2-\delta}, \frac{1-\delta}{2-\delta}\right)$ . No further information will be provided subsequently. The optimal action to play is  $d$  and the agent's conditional payoff is 0.

Overall the agent's expected payoff for selecting option 2 is *at most*

$$\underbrace{\frac{2-\delta}{3}}_{\mathbb{P}(g_{\alpha,\gamma} | S_2)} \times \frac{1}{1+\delta} \quad + \quad \underbrace{\frac{1+\delta}{3}}_{\mathbb{P}(g_{\beta,\gamma} | S_2)} \times 0 \quad = \quad \frac{2-\delta}{3(1+\delta)},$$

which is smaller than the payoff obtained by selecting option 1. Hence, a type- $t_1$  agent would strictly prefer option 1.

By symmetry, the same analysis can be applied to show that a type- $t_2$  agent would strictly prefer option 2. This completes the proof.  $\square$

### A.3 Proof of Lemma 5

**Lemma 5.** *Suppose that  $\mathbb{E}(v) \geq 1 - \delta$  in Example 4.2. Then, for any  $\epsilon \in (0, 1 - \delta)$ , there exists a signal  $g \in G$  such that  $\mathbb{E}(v | g) > 1 - \delta - \epsilon$  and  $\mathbb{P}(\gamma | g) \geq 1/4 - \delta/\epsilon$ .*

*Proof.* Suppose for a contradiction that for every  $g \in G$ , at least one of  $\mathbb{E}(v | g) > 1 - \delta - \epsilon$  and  $\mathbb{P}(\gamma | g) \geq 1/4 - \delta/\epsilon$  is violated. Let  $G = G^- \cup G^+$ , where

$$G^- = \{g \in G : \mathbb{P}(\gamma | g) < 1/4 - \delta/\epsilon\} \quad \text{and} \quad G^+ = \{g \in G : \mathbb{P}(\gamma | g) \geq 1/4 - \delta/\epsilon\}.$$

Note that by definition,

$$\sum_{g \in G} \mathbb{P}(g) \cdot \mathbb{P}(\gamma | g) = \mathbb{P}(\gamma) = 1/4.$$

We have

$$\begin{aligned}
1/4 &= \sum_{g \in G^-} \mathbb{P}(g) \cdot \mathbb{P}(\gamma | g) + \sum_{g \in G^+} \mathbb{P}(g) \cdot \mathbb{P}(\gamma | g) \\
&< (1/4 - \delta/\epsilon) \cdot \sum_{g \in G^-} \mathbb{P}(g) + \sum_{g \in G^+} \mathbb{P}(g) = (1/4 - \delta/\epsilon) \cdot (1 - \mathbb{P}(G^+)) + \mathbb{P}(G^+),
\end{aligned}$$

where we write  $\mathbb{P}(G^+) = \sum_{g \in G^+} \mathbb{P}(g)$ . It follows that

$$\mathbb{P}(G^+) > \frac{\delta/\epsilon}{3/4 + \delta/\epsilon} > \frac{\delta}{\epsilon + \delta}. \quad (18)$$

According to the assumption at the beginning of this proof,  $\mathbb{E}(v | g) \leq 1 - \delta - \epsilon$  for all  $g \in G^+$ . Consider the expected payoff  $\mathbb{E}(v)$  the mechanism yields for the principal. We have

$$\begin{aligned}
\mathbb{E}(v) &= \sum_{g \in G^+} \mathbb{P}(g) \cdot \mathbb{E}(v | g) + \sum_{g \in G^-} \mathbb{P}(g) \cdot \mathbb{E}(v | g) \\
&< (1 - \delta - \epsilon) \cdot \sum_{g \in G^+} \mathbb{P}(g) + \sum_{g \in G^-} \mathbb{P}(g) \cdot 1 \\
&= (1 - \delta - \epsilon) \cdot \mathbb{P}(G^+) + (1 - \mathbb{P}(G^+)) \\
&= 1 - (\delta + \epsilon) \cdot \mathbb{P}(G^+) \\
&< 1 - \delta,
\end{aligned}$$

where we use the fact that the principal's maximum attainable payoff is 1 in Example 4.2 to bound  $\mathbb{E}(v | g)$  for  $g \in G^-$ , and the last transition follows by (18). This contradicts the assumption that  $\mathbb{E}(v) \geq 1 - \delta$ .  $\square$

#### A.4 Proof of Lemma 6

**Lemma 6.** *Suppose that  $\mathbb{E}(v) \geq 1 - \delta$  in Example 4.2. Then there exists a signal  $g \in G$  such that  $\mathbb{E}(v | g) > 1 - 11\delta$ ,  $\mathbb{P}(\alpha | g) > 1/100$ , and  $\mathbb{P}(\beta | g) > 1/100$ .*

*Proof.* Given Lemma 5, we can find a signal  $g$  such that

$$\mathbb{E}(v | g) > 1 - \delta - \epsilon \quad (19)$$

$$\text{and } \mathbb{P}(\gamma | g) \geq 1/4 - \delta/\epsilon. \quad (20)$$

Specifically, choosing  $\epsilon = 10\delta$  (recall that we have  $M > 10000$  and  $\delta = 10/M$ , so  $\epsilon < 1$ ), we get that

$$\mathbb{E}(v | g) > 1 - 11\delta \quad (21)$$

$$\text{and } \mathbb{P}(\gamma | g) > 1/10. \quad (22)$$

First, note that since the principal gets payoff 1 only when the agent plays  $a'$  or  $b'$ , and 0 otherwise, it must be that

$$\mathbb{P}[t_1, \gamma, \neg(a' \vee b') \mid g] < \delta + \epsilon,$$

i.e., the probability that a type- $t_1$  agent plays an action outside of  $\{a', b'\}$  when the state is  $\gamma$  cannot exceed  $\delta + \epsilon$ . Indeed, if this inequality does not hold, we would have  $\mathbb{E}(v | g) \leq 1 - \delta - \epsilon = 1 - 11\delta$ , contradicting (19). It follows that

$$\mathbb{P}(t_1 | g) \cdot \mathbb{P}[\gamma, \neg(a' \vee b') \mid t_1, g] < \delta + \epsilon \quad (23)$$

Since the occurrence of type  $t_1$  is independent of the occurrence of  $g$ , we have  $\mathbb{P}(t_1 | g) = \mathbb{P}(t_1) = \rho(t_1) = 1/3$ . Hence,

$$\mathbb{P}[\gamma, \neg(a' \vee b') \mid t_1, g] < 3(\delta + \epsilon),$$

and in turn

$$\begin{aligned} \mathbb{P}(\gamma, a' \vee b' \mid t_1, g) &> \mathbb{P}(\gamma \mid t_1, g) - 3(\delta + \epsilon) \\ &> 1/10 - 3(\delta + \epsilon), \end{aligned} \tag{24}$$

where we used the fact that  $\mathbb{P}(\gamma \mid t_1, g) = \mathbb{P}(\gamma \mid g)$  due to the independence between  $\gamma$  and  $t_1$  conditioned on  $g$ .

We now turn to the following upper bound of a type- $t_1$  agent's expected payoff, conditioned on them receiving  $g$  at the first step:

$$\mathbb{E}[u_{t_1} \mid t_1, g] \leq \mathbb{P}(\gamma, b' \mid t_1, g) \cdot (-M) + (1 - \mathbb{P}(\gamma, b' \mid t_1, g)) \cdot 3,$$

where  $-M = u_{t_1}(b', \gamma) = u_{t_1}(c', \gamma)$  and 3 is the maximum attainable payoff for any possible combination of action and state, according to the agent's payoffs in Example 4.2. It must be that  $\mathbb{E}[u_{t_1} \mid t_1, g] \geq 0$  because otherwise the type- $t_1$  agent would be strictly better off if they stick to action  $d$  after receiving  $g$ , whereby a payoff of 0 is guaranteed. Hence, we have

$$\mathbb{P}(\gamma, b' \mid t_1, g) \cdot (-M) + (1 - \mathbb{P}(\gamma, b' \mid t_1, g)) \cdot 3 \geq 0,$$

which implies

$$\mathbb{P}(\gamma, b' \mid t_1, g) \leq \frac{3}{M+1}.$$

Since playing  $a'$  and playing  $b'$  are mutually exclusive events, it follows by this inequality and (24) that

$$\mathbb{P}(\gamma, a' \mid t_1, g) > \frac{1}{10} - 3(\delta + \epsilon) - \frac{3}{M+1}. \tag{25}$$

Similarly to the upper bound of  $\mathbb{E}[u_{t_1} \mid t_1, g]$ , the following upper bound can be established for a type- $t_1$  agent's expected payoff  $\mathbb{E}[u_{t_1} \mid a', t_1, g]$  conditioned on  $g$  and  $a'$  being played:

$$\begin{aligned} \mathbb{E}[u_{t_1} \mid a', t_1, g] &= \mathbb{P}(\alpha \mid a', t_1, g) \cdot 1 + \mathbb{P}(\gamma \mid a', t_1, g) \cdot (-1) + \mathbb{P}(\beta \mid a', t_1, g) \cdot (-M) \\ &\leq \mathbb{P}(\alpha \mid a', t_1, g) - \mathbb{P}(\gamma \mid a', t_1, g). \end{aligned}$$

According to (25),  $\mathbb{P}(\gamma, a' \mid t_1, g) > 0$ , which means action  $a'$  is indeed played by a type- $t_1$  agent with a positive probability conditioned on  $g$ . It must then be that  $\mathbb{E}[u_{t_1} \mid a', t_1, g] \geq 0$ : if the type- $t_1$  agent were to expect a negative payoff for playing  $a'$ , then they would be better off deviating from  $a'$  to  $d$ . Note in particular that when  $a'$  is played, it must be that the type- $t_1$  agent reported their type truthfully because  $a'$  is strictly dominated by  $a$  or  $d$  for types  $t_2$  for any posterior distribution over the states— $a'$  would not be played if  $t_2$  were reported. Hence, the above described deviation will not be blocked by the credibility constraint. As a result,  $\mathbb{P}(\alpha \mid a', t_1, g) \geq \mathbb{P}(\gamma \mid a', t_1, g)$ , which

implies  $\mathbb{P}(\alpha, a' | t_1, g) \geq \mathbb{P}(\gamma, a' | t_1, g)$  by multiplying the two sides with  $\mathbb{P}(a' | t_1, g)$ . It follows that

$$\begin{aligned}
\mathbb{P}(\alpha | g) &\geq \mathbb{P}(t_1 | g) \cdot \mathbb{P}(\alpha | t_1, g) \\
&= \rho(t_1) \cdot \mathbb{P}(\alpha | t_1, g) \\
&\geq \rho(t_1) \cdot \mathbb{P}(\alpha, a' | t_1, g) \\
&\geq \rho(t_1) \cdot \mathbb{P}(\gamma, a' | t_1, g) \\
&\geq \frac{1}{30} - (\delta + \epsilon) - \frac{2}{3(M+1)} && \text{(by (25))} \\
&> \frac{1}{100}. && \text{(as } M > 10000, \delta = 10/M, \text{ and } \epsilon = 10\delta)
\end{aligned}$$

By symmetry the same argument also applies to deriving:  $\mathbb{P}(\beta | g) > 1/100$  by changing  $t_1$  to  $t_2$  and  $\alpha$  to  $\beta$  in the proof.  $\square$

## A.5 Proof of Theorem 9

**Theorem 9.** *For any indefinite-stage mechanism  $\Pi$ , there exists an  $\tilde{\text{ESES}}$  mechanism  $\Pi'$  that yields as much payoff for the principal as  $\Pi$  does and has the following properties.*

- i. **IC at  $\tilde{\text{E}}$ :** *The root  $\tilde{\text{E}}$  has  $n$  children  $S_1, \dots, S_n$ , which are all  $S$  nodes. Each type- $t_i$  agent is incentivized to choose the branch leading to  $S_i$ .*
- ii. **DIC at  $S_i$ :** *Each node  $S_i$  has  $n$   $E$  node children  $E_{i1}, \dots, E_{in}$ . Each  $E_{ij}$  is associated with type  $t_j$ , and at  $E_{ij}$  a type- $t_i$  agent is incentivized to report  $t_j$ . For each  $k = 1, \dots, n$ , reporting type  $t_k$  at  $E_{ij}$  leads to an  $S$  node child  $S_{ijk}$ .*
- iii. **DIC at  $S_{ijk}$ :** *The signaling strategy used at each node  $S_{ijk}$  directly recommends the agent an action to take, and the action is optimal for type  $t_k$ .*

*Proof.* Given any  $\Pi$ , we will construct an  $\tilde{\text{ESES}}$  mechanism  $\Pi'$  and show that the conversion ensures the stated properties while preserving the principal's payoff. Without loss of generality, we assume that there is exactly one  $E$  node on every path from the root to a leaf. Indeed, there is no need to elicit the agent's type more than once due to the credibility requirement. On the other hand, if there is no  $E$  on some path, we can replace the leaf node on this path with an  $E$  node (whose children are all leaf nodes). Moreover, we assume that every path begins and ends with signaling stages. This can be achieved by adding a dummy  $S$  node as the parent of the root node of  $\Pi$  where the same signal is sent irrespective of the states, and similarly by inserting a dummy  $S$  node before every leaf node.

Next, we apply the following operations to construct  $\Pi'$ .

**Step 1. Adding a New root** First, we create an  $\tilde{\text{E}}$  node as the root node of  $\Pi'$ . Then we make  $n$  copies  $\Pi^1, \dots, \Pi^n$  of  $\Pi$  and attach each  $\Pi^i$  to the root of  $\Pi'$  as a subtree. We mark the root of each  $\Pi^i$  as  $S_i$ . It is straightforward that  $\Pi'$  constructed so far is IC at the root node: it is optimal for each type  $t_i$  to select the  $i$ -th child  $\Pi^i$  of  $\Pi'$ . Since all the copies are the same, each type  $t_i$  is incentivized to choose  $\Pi^i$ , so  $\Pi'$  is IC at the root. Moreover, the payoffs  $\Pi'$  yields for both the principal and every agent type are the same as those yielded by  $\Pi$ .



**Step 2. Removing  $\tilde{\text{E}}$  nodes** Given IC at the root, in each branch  $\Pi^i$ , we only need to consider the response of  $t_i$ . Hence, all the  $\tilde{\text{E}}$  nodes in  $\Pi^i$  become redundant. For each  $\tilde{\text{E}}$  node  $e$  in  $\Pi^i$ , we remove  $e$ , connecting the branch of  $e$  a type- $t_i$  agent is incentivized to select directly to the parent of  $e$ . (All other branches of  $e$  are removed.) It is straightforward that this operation preserves the payoff of each type  $t_i$  for  $\Pi^i$ , while it can only decrease their payoff for each  $\Pi^j$ ,  $j \neq i$ , now that there are fewer choices in each  $\Pi^j$ . Hence,  $\Pi'$  remains IC and yields the same payoffs for the players as before.

**Step 3. Merging S nodes** Next, we merge consecutive S nodes on every path. If any S node  $s$  has a child  $s'$  which is also an S node, we remove the edge connecting  $s$  and  $s'$  and attach the children of  $s'$  directly to  $s$ . We also modify the signaling strategy  $\pi_s$  at  $s$ : replace the signal  $g$  in  $\pi_s$  that corresponds to  $s'$  with signals in the support of  $\pi_{s'}$ , and redefine  $\pi_s(g | \theta) = 0$  and

$$\pi_s(g' | \theta) = \pi_{s'}(g | \theta) \cdot \pi_{s'}(g' | \theta)$$

for each signal  $g'$  in the support of  $\pi_{s'}$ . Namely, the signals sent at  $s'$  is now generated directly at  $s$ . The modified probabilities in  $\pi_s$  ensures that the posterior distributions induced by these signals remain the same as before.

The above operations yield a five-level tree. Levels 1 to 5 contain only  $\tilde{\text{E}}$ , S, E, S, and leaf nodes, respectively. The mechanism is IC at  $\tilde{\text{E}}$ , but may not yet be DIC at the other S nodes. We proceed with the following steps to ensure the DIC properties.

**Step 4. Ensuring DIC at  $S_i$  and  $S_{ijk}$**  Recall that  $S_i$  is the root of  $\Pi^i$  and now at the second level of  $\Pi'$ , and  $S_{ijk}$  is at the fourth level. To ensure DIC, we merge signals sent at these nodes that induce the same agent response (i.e., reporting the same type or taking the same action).

Specifically, consider an arbitrary  $S_i$  and let us first identify the following sets of nodes:

- For each  $j = 1, \dots, n$ , let  $\mathcal{E}_{ij}$  be the set consisting of the children of  $S_i$  (which are all E nodes) at which a type- $t_i$  agent is incentivized to report  $t_j$ .
- For each  $k = 1, \dots, n$ , we let  $\mathcal{S}_{ijk}$  be the set consisting of the children of  $\mathcal{E}_{ij}$  that correspond to the agent reporting type  $t_k$ .

We merge nodes in each  $\mathcal{E}_{ij}$  into a single E node. In other words, we replace signals leading to these nodes with a single signal, which can be viewed as a signal that recommends the agent to report  $t_j$ . This operation also makes the nodes in each  $\mathcal{S}_{ijk}$  indistinguishable for the agent. In other words, these nodes are now in the same information set. We treat each information set as a new node and denote them by  $\text{E}_{ij}$  or  $\text{S}_{ijk}$ , accordingly. The signaling strategies at  $S_i$  and  $S_{ijk}$  are updated accordingly as follows.

- We update the strategy  $\pi_i$  at  $S_i$  by merging the probabilities of the merged signals: for every  $\theta \in \Theta$ ,

$$\pi_i(g_{\text{E}_{ij}} | \theta) \leftarrow \sum_{e \in \mathcal{E}_{ij}} \pi_i(g_e | \theta)$$

where  $g_e$  denotes the signal leading to a node  $e$  (and we abandon all the original signals  $g_e$  for all  $e \in \mathcal{E}_{ij}$ ).

- The signaling strategy  $\pi_{ijk}$  at each  $S_{ijk}$  is defined over  $\bigcup_{e \in S_{ijk}} G_e$ , where  $G_e$  denotes the signal space of the strategy at  $e$ . (W.l.o.g., the  $G_e$ 's are disjoint sets because in  $\Pi$  the agent always knows which node  $e$  they are at.) For every  $g \in \bigcup_{e \in S_{ijk}} G_e$ , we let

$$\pi_{ijk}(g | \theta) \leftarrow \sum_{e \in \mathcal{E}_{ij}} w_e \cdot \pi_{ek}(g | \theta),$$

where  $ek$  denotes the child of  $e$  corresponding to type  $t_k$ , and  $\pi_{ek}$  is the signaling strategy at  $ek$ ; the normalization factor  $w_e := \frac{\pi_i(g_e)}{\sum_{e' \in \mathcal{E}_{ij}} \pi_i(g_{e'})}$  is the probability of  $e$  conditioned on  $\mathcal{E}_{ij}$ .

The above design ensures that each  $g$  sent by  $\pi_{ijk}$  induces the same posterior belief as it previously does. Hence, at the new node  $E_{ij}$ , a type- $t_i$  agent would still be incentivized to report  $t_j$ , so DIC is ensured at  $S_i$ .

DIC at each  $S_{ijk}$  can be ensured similarly by merging signals that induce the same action of the agent. Essentially, the modifications in Step 4 make many nodes in  $\Pi^i$  indistinguishable to the agent, but if the agent follows the recommendations in the new mechanism, they would still reach each leaf with the same probability as before, as if they respond optimally in  $\Pi^i$ . On the other hand, since these merging operations reduce information, the agent cannot devise any better response to improve their payoff. Hence, the agent is incentivized to follow the recommendations and this ensures DIC.  $\square$

## A.6 Proof of Lemma 13

**Lemma 13.**  $|\Theta_x| = 1$  for every leaf node  $x$  in  $\Pi$ .

*Proof.* Since  $\Pi$  is optimal, it yields payoff 1 for the principal. This means that only action  $a'_i$  can be played in state  $\theta_i$ , so no action will be played at two different states both with positive probabilities. Now that  $x$  is a leaf node, the agent needs to choose an action to play. However, since  $\Theta_x$  contains more than one state, no matter which action the agent chooses, the same action will be played at more than one state, a contradiction.  $\square$

## A.7 Proof of Lemma 14

**Lemma 14.** For every node  $x$  in  $\Pi$ , if  $\theta_i \in \Theta_x$  then  $t_i \in T_x$ .

*Proof.* Suppose that  $\theta_i \in \Theta_x$  but  $t_i \notin T_x$ . Since  $\Pi$  is optimal, it yields payoff 1 for the principal. This means that action  $a'_i$  must be played in state  $\theta_i$ , so we have  $\mathbb{P}(a'_i | x) > 0$ . Nevertheless,  $a'_i$  is strictly dominated for all types in  $T \setminus \{t_i\}$ , so it will never be played if the agent imitates any type in  $T_x$ , now that  $t_i \notin T_x$ . Hence, there is a contradiction.  $\square$

## A.8 Proof of Lemma 15

We first prove the following lemma.

**Lemma 18.** For every node  $x$  in  $\Pi$ , if  $t_i \in T_x$ , then  $p_x(\theta_j) \leq p_x(\theta_n) + 2/M$  for all  $j = i + 1, \dots, n$ .

*Proof.* Suppose for the sake of contradiction that for some  $j \in \{i + 1, \dots, n\}$  it holds that

$$p_x(\theta_j) > p_x(\theta_n) + 2/M \tag{26}$$

Since  $\Pi$  is optimal, according to our analysis in Section 6.2.1, it must yield payoff 1 for the principal. Hence, only the outcomes  $(\theta_i, a'_i)$ ,  $i = 1, \dots, n$ , have positive probabilities in the probability measure

induced by  $\Pi$ . In this case, a type- $t_0$  agent's expected payoff conditioned on  $x$  is  $\sum_{i=1}^n p_x(\theta_i) \cdot u_{t_0}(\theta_i, a'_i) = p_x(\theta_n)$ . We demonstrate the following contradiction: there is a way for a type- $t_0$  agent to obtain strictly more than  $p_x(\theta_n)$ .

Consider the outcomes induced if a type- $t_0$  agent imitates to  $t_i$  in the subtree starting at  $x$ ; this is feasible given that  $t_i \in T_x$ . Namely, the agent always selects a subset containing  $t_i$  in response to elicitation in the subtree and eventually plays an optimal action of  $t_i$ . Let  $\mathbb{P}_*$  be the probability measure induced by  $\Pi$  and this strategy of the agent. According to the incentive of type  $t_i$ , if  $\mathbb{P}_*(a'_i | x) > 0$ , then it must be that

$$\mathbb{P}_*(-\theta_i | a'_i, x) < 1/M^2. \quad (27)$$

Indeed, if  $\mathbb{P}_*(-\theta_i | a'_i, x) \geq 1/M^2$ , type  $t_i$ 's expected payoff conditioned on  $a'_i$  being played in the subtree would be *at most*

$$\mathbb{P}_*(-\theta_i | a'_i, x) \cdot (-M) + (1 - \mathbb{P}_*(-\theta_i | a'_i, x)) \cdot \underbrace{1/M}_{u_{t_i}(\theta_i, a'_i)} < 0.$$

In this case, type  $t_i$  would be overall better off swapping action  $a'_i$  with  $a_i$  whenever the former is to be played, contradicting the assumption that  $\mathbb{P}_*(a'_i | x) > 0$ .

Now consider the type- $t_0$  agent's actual payoff over  $\mathbb{P}_*$ , that is  $\mathbb{E}_*(u_{t_0} | x) = \sum_{\theta \in \Theta} \sum_{a \in A} \mathbb{P}_*(\theta, a | x) \cdot u_{t_0}(\theta, a)$ . We can exclude actions in  $\{a'_1, \dots, a'_n\} \setminus \{a'_i\}$ , which are strictly dominated by other actions according to type  $t_i$ 's incentive. It follows that

$$\begin{aligned} \mathbb{E}_*(u_{t_0} | x) &= \sum_{\theta \in \Theta} \sum_{k=1}^n \mathbb{P}_*(\theta, a_k | x) \cdot u_{t_0}(\theta, a_k) + \sum_{\theta \in \Theta} \mathbb{P}_*(\theta, a'_i | x) \cdot u_{t_0}(\theta, a'_i) \\ &\geq \sum_{\theta \in \Theta} \sum_{k=1}^n \mathbb{P}_*(\theta, a_k | x) \cdot \underbrace{u_{t_i}(\theta, a_k)}_{\text{as } u_{t_0}(\theta, a_k) \geq u_{t_i}(\theta, a_k)} + \underbrace{\mathbb{P}_*(-\theta_i, a'_i | x) \cdot (-M)}_{\text{as } u_{t_0}(\theta_i, a'_i) = 0 \text{ and } u_{t_0}(\theta, a'_i) = -M \forall \theta \neq \theta_i} \end{aligned} \quad (28)$$

The second term in the last line can be bounded from below by  $-1/M$  because we have

$$\mathbb{P}_*(-\theta_i, a'_i | x) = \mathbb{P}_*(a'_i | x) \cdot \mathbb{P}_*(-\theta_i | a'_i, x) \leq \mathbb{P}_*(-\theta_i | a'_i, x) < 1/M^2$$

where  $\mathbb{P}_*(-\theta_i | a'_i, x) < 1/M^2$  is implied by (27). We next bound the first term. By definition,  $\mathbb{P}_*$  is induced by a type- $t_i$  agent's best response. So, the first term can be bounded from below by what a type- $t_i$  agent would get if he swapped all the actions  $a_1, \dots, a_n$  to  $a_j$ . (Recall  $j$  is the number that makes (26) hold.) Specifically, we have

$$\begin{aligned} \sum_{\theta \in \Theta} \sum_{k=1}^n \mathbb{P}_*(\theta, a_k | x) \cdot u_{t_i}(\theta, a_j) &= \sum_{k=1}^n \mathbb{P}_*(\theta_j, a_k | x) \cdot \underbrace{u_{t_i}(\theta_j, a_j)}_{=1} + \sum_{\theta \neq \theta_j} \sum_{k=1}^n \mathbb{P}_*(\theta, a_k | x) \cdot \underbrace{u_{t_i}(\theta, a_j)}_{=0} \\ &= \sum_{k=1}^n \mathbb{P}_*(\theta_j, a_k | x) \\ &= \mathbb{P}_*(\theta_j | x) - \mathbb{P}_*(\theta_j, a'_i | x) - \sum_{k \neq i} \underbrace{\mathbb{P}_*(\theta_j, a'_k | x)}_{=0 \text{ as } a'_k \text{ is strictly dominated}} \\ &= \mathbb{P}_*(\theta_j | x) - \mathbb{P}_*(\theta_j, a'_i | x) \\ &\geq \mathbb{P}_*(\theta_j | x) - \mathbb{P}_*(\theta_j | a'_i, x) > \mathbb{P}_*(\theta_n | x) + 2/M - 1/M^2, \end{aligned}$$

where the last transition is due to (26) and (27). Substituting into (28) this lower bound and the bound  $-1/M$  we derived for the second term, we get the contradiction  $\mathbb{E}_*(u_{t_0} | x) > p_x(\theta_n)$ , as desired.  $\square$

**Lemma 15.** *For every  $i = 1, \dots, n-1$  and  $\epsilon \in \mathbb{R}$ , let  $\mathcal{P}_i^\epsilon \subseteq \Delta(\Theta)$  consist of points  $p$  such that:*

- i.  $\sum_{j=i}^n p(\theta_j) = 1$ ;
- ii.  $\sum_{j=i+1}^n p(\theta_j) \geq 1 - \frac{1}{n-i+1} - \epsilon$ ; and
- iii.  $|p(\theta_j) - p(\theta_n)| \leq \epsilon/2$  for all  $j = i+1, \dots, n$ .

Let  $\epsilon_k = n^{-3^{n-k}}$  for  $k \in \mathbb{Z}$ . For every non-leaf node  $x$  in  $\Pi$ , if  $t_i \in T_x$  and  $p_x \in \mathcal{P}_i^{\epsilon_k}$ , then there exists  $c \in \text{children}(x)$  such that  $p_c \in \mathcal{P}_i^{\epsilon_{k+1}}$ .

*Proof.* If  $x$  is an E node, the posteriors at its children would be the same as that at  $x$ . The statement holds trivially. Hence, in what follows we consider the case where  $x$  is an S node.

Suppose for the sake of contradiction that  $t_i \in T_x$  and  $p_x \in \mathcal{P}_i^{\epsilon_k}$ , but  $p_c \notin \mathcal{P}_i^{\epsilon_{k+1}}$  for all  $c \in \text{children}(x)$ . Note that condition (i) holds automatically for all  $c \in \text{children}(x)$  because  $p_c(\theta_{j'})$  remains 0 for all  $j < i$ . Hence, every  $c \in \text{children}(x)$  either violates (ii)  $\sum_{j=i+1}^n p_c(\theta_j) \geq 1 - \frac{1}{n-i+1} - \epsilon_{k+1}$ , or (iii)  $|p_c(\theta_j) - p_c(\theta_n)| \leq \epsilon_{k+1}/2$  for some  $j$ .

We can then define a partition  $\{C_0, C_1, \dots, C_n\}$  of  $\text{children}(x)$ , where  $C_0$  contains children violating (ii), and each other  $C_j$  contains children violating (iii) for  $j$ . (If a child violates multiple conditions, we assign it to an arbitrary set.)

We next bound the conditional probabilities  $\mathbb{P}(C_j | x) = \sum_{c \in C_j} \mathbb{P}(c | x)$  of these subsets, where  $\mathbb{P}$  is the probability measure induced by  $\Pi$ . Note that for every  $j$  we can write  $p_x(\theta_j)$  as the weighted average  $p_x(\theta_j) = \sum_{c \in \text{children}(x)} \mathbb{P}(c | x) \cdot p_c(\theta_j)$ .

- Every  $c \in C_0$  violates the first condition, so  $\sum_{j=i+1}^n p_c(\theta_j) < 1 - \frac{1}{n-i+1} - \epsilon_{k+1}$ . It must be that

$$\mathbb{P}(C_0 | x) \leq \xi := \frac{\frac{1}{n-i+1} + \epsilon_k}{\frac{1}{n-i+1} + \epsilon_{k+1}}. \quad (29)$$

Indeed, if  $\mathbb{P}(C_0 | x) > \xi$  was true, we would have

$$\begin{aligned} \sum_{j=i+1}^n p_x(\theta_j) &= \sum_{j=i+1}^n \left( \sum_{c \in C_0} \mathbb{P}(c | x) \cdot p_c(\theta_j) + \sum_{c \in \text{children}(x) \setminus C_0} \mathbb{P}(c | x) \cdot p_c(\theta_j) \right) \\ &= \sum_{c \in C_0} \mathbb{P}(c | x) \sum_{j=i+1}^n p_c(\theta_j) + \sum_{c \in \text{children}(x) \setminus C_0} \mathbb{P}(c | x) \sum_{j=i+1}^n p_c(\theta_j) \\ &< \mathbb{P}(C_0 | x) \cdot \left( 1 - \frac{1}{n-i+1} - \epsilon_{k+1} \right) + (1 - \mathbb{P}(C_0 | x)) \cdot 1 \\ &\leq \xi \cdot \left( 1 - \frac{1}{n-i+1} - \epsilon_{k+1} \right) + (1 - \xi) \quad (\text{assuming } \mathbb{P}(C_0 | x) > \xi) \\ &= 1 - \frac{1}{n-i+1} - \epsilon_k, \end{aligned}$$

contradicting the assumption that  $p_x \in \mathcal{P}_i^{\epsilon_k}$ .

- For every  $j = i+1, \dots, n$ , every  $c \in C_j$  violates the second condition for  $j$ , so  $|p_c(\theta_j) - p_c(\theta_n)| > \epsilon_{k+1}/2$ . By assumption  $t_i \in T_x$ . Now that  $x$  is an S node, no type will be excluded from  $T_x$  at  $x$ , so we have  $T_c = T_x$  for all  $c \in \text{children}(x)$ . It then follows by Lemma 18 that:

$$p_c(\theta_j) \leq p_c(\theta_n) + 2/M. \quad (30)$$

Since  $2/M < \epsilon_{k+1}/2$  is sufficiently small,  $|p_c(\theta_j) - p_c(\theta_n)| > \epsilon_{k+1}/2$  holds only if

$$p_c(\theta_j) < p_c(\theta_n) - \epsilon_{k+1}/2. \quad (31)$$

Then, it must be that

$$\mathbb{P}(C_j | x) \leq \xi' := 2\epsilon_k/\epsilon_{k+1}, \quad (32)$$

as the converse,  $\mathbb{P}(C_j | x) > \xi'$ , would result in

$$\begin{aligned} p_x(\theta_j) &= \sum_{c \in C_j} \mathbb{P}(c | x) \cdot p_c(\theta_j) + \sum_{c \in \text{children}(x) \setminus C_j} \mathbb{P}(c | x) \cdot p_c(\theta_j) \\ &\leq \sum_{c \in C_j} \mathbb{P}(c | x) \cdot \underbrace{(p_c(\theta_n) - \epsilon_{k+1}/2)}_{\text{by (31)}} + \sum_{c \in \text{children}(x) \setminus C_j} \mathbb{P}(c | x) \cdot \underbrace{(p_c(\theta_n) + 2/M)}_{\text{by (30)}} \\ &= \sum_{c \in \text{children}(x)} \mathbb{P}(c | x) \cdot p_c(\theta_n) - \mathbb{P}(C_j | x) \cdot \epsilon_{k+1}/2 + (1 - \mathbb{P}(C_j | x)) \cdot 2/M \\ &< p_x(\theta_n) - \epsilon_k + 2/M \\ &< p_x(\theta_n) - \epsilon_k/2, \end{aligned}$$

contradicting  $p_x \in \mathcal{P}_i^{\epsilon_k}$ , too.

Summing up (29) and (32) gives the following contradiction, for  $n$  sufficiently large:

$$\begin{aligned} \sum_{j=0}^n \mathbb{P}(C_j | x) &\leq \xi + n\xi' = \frac{\frac{1}{n-i+1} + \epsilon_k}{\frac{1}{n-i+1} + \epsilon_{k+1}} + n \cdot 2\epsilon_k/\epsilon_{k+1} \\ &\leq \frac{1 + \epsilon_k}{1 + \epsilon_{k+1}} + n \cdot 2\epsilon_k/\epsilon_{k+1} \leq \frac{\epsilon_{k+1} + n^2 \cdot \epsilon_k}{\epsilon_{k+1} + \epsilon_{k+1}^2} < 1, \end{aligned}$$

where we used the fact that  $\epsilon_{k+1}^2 > n^2 \epsilon_k$  (recall that by definition  $\epsilon_k = n^{-3^{n-k}}$  for  $k \in \mathbb{Z}$ ).  $\square$