

Partial Truthfulness in Minimal Peer Prediction Mechanisms with Limited Knowledge

Goran Radanovic

Harvard University
Cambridge, USA
gradanovic@g.harvard.edu

Boi Faltings

EPFL
Lausanne, Switzerland
boi.faltings@epfl.ch

Abstract

We study minimal single-task peer prediction mechanisms that have limited knowledge about agents' beliefs. Without knowing what agents' beliefs are or eliciting additional information, it is not possible to design a truthful mechanism in a Bayesian-Nash sense. We go beyond truthfulness and explore equilibrium strategy profiles that are only partially truthful. Using the results from the multi-armed bandit literature, we give a characterization of how inefficient these equilibria are comparing to truthful reporting. We measure the inefficiency of such strategies by counting the number of dishonest reports that any minimal knowledge-bounded mechanism must have. We show that the order of this number is $\Theta(\log n)$, where n is the number of agents, and we provide a peer prediction mechanism that achieves this bound in expectation.

Introduction

One of the crucial prerequisites for a good decision making procedure is the availability of accurate information, which is often distributed among many individuals. Hence, elicitation of distributed information represents a key component in many systems that are based on informative decision making. Typically, such an information elicitation scenario is modeled by representing individuals as rational agents who are willing to report their private information in return for (monetary) rewards.

We study a setting in which reports cannot be directly verified, as it is the case when eliciting opinions regarding the outcome of a hypothetical question (Garcin and Faltings 2014). Other examples include: product reviewing, where the reported information can describe ones taste, which is inherently subjective¹; peer grading, where a rater needs to grade an essay; or eliciting information that is highly distributed, as in community sensing.

Since a data collector cannot directly verify the reported information, it can score submitted reports only by examining consistency among them. Such an approach is adopted in *peer prediction mechanisms*, out of which the most known examples are the *peer prediction method* (Miller, Resnick,

and Zeckhauser 2005) and the *Bayesian truth serum* (Prelec 2004). While there are different ways of classifying peer prediction mechanisms, the most relevant one for this work distinguishes two categories of mechanisms by: 1) the amount of additional information they elicit from agents; 2) the knowledge they have about agents' beliefs.

The first category includes *minimal* mechanisms that elicit only desired private information, but have some additional information that enables truthful elicitation. For instance, the classical peer prediction (Miller, Resnick, and Zeckhauser 2005) assumes knowledge about how an agent forms her beliefs regarding the private information of other agents, while other mechanisms (e.g., see (Jurca and Faltings 2011; Witkowski 2014)) relax the amount of knowledge they need by imposing different restrictions on the agents' belief structures. Mechanisms from the second category elicit additional information to compensate for the lack of knowledge about agents' beliefs. For example, the Bayesian truth serum (Prelec 2004), and its extensions (Witkowski and Parkes 2012b; Radanovic and Faltings 2013; Radanovic and Faltings 2014; Kong and Schoenebeck 2016), elicit agents' posterior beliefs.

The mentioned mechanisms are designed for a *single-task* elicitation scenario in which an agent's private information can be modeled as a sample from an unknown distribution. In a basic elicitation setting, agents share a common belief system regarding the parameters of the setting (Miller, Resnick, and Zeckhauser 2005; Prelec 2004). While the mechanisms typically allow some deviations from this assumption (Frongillo and Witkowski 2016; Radanovic and Faltings 2014), these deviations can be quite constrained, especially when private information has a complex structure.²

We also mention mechanisms that operate in more specialized settings that allow agents to have more heterogeneous beliefs. Peer prediction without a common prior (Witkowski and Parkes 2012a) is designed for a setting in which a mechanism can clearly separate a period prior to agents acquiring their private information from the period after the acquisition. It elicits additional information from agents which corresponds to their prior beliefs. More

Copyright by the authors. This is an extended version of the paper that will appear in: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI'18)*.

¹As, for example, in rating a restaurant or a hotel on TripAdvisor (www.tripadvisor.com).

²For example, (Radanovic and Faltings 2014) show that one cannot easily relax the common prior condition when agents' private information is real-valued.

recently, many mechanisms have been developed for a multi-task elicitation scenario (Dasgupta and Ghosh 2013; Kamble et al. 2015; Radanovic, Faltings, and Jurca 2016; Shnayder et al. 2016a), primarily designed for crowdsourcing settings. The particularities of the multi-task setting enable the mechanisms to implicitly extract relevant information important for scoring agents (e.g., agents’ prior beliefs). For more details on peer prediction mechanisms, we refer the reader to (Faltings and Radanovic 2017).

Clearly, there is a tradeoff between the assumed knowledge and the amount of elicited information. The inevitability of such a trade-off can be expressed by a result of (Radanovic and Faltings 2013), which states that in the single-task elicitation setting, no minimal mechanism can achieve truthfulness for a general common belief system.

Contributions. In this paper, we investigate single-task *minimal* peer prediction mechanisms that have only *limited* information about the agents’ beliefs, which precludes them from incentivizing all agents to report honestly. To characterize the inefficiency of such an approach, we introduce a concept of *dishonesty limit* that measures the minimal number of dishonest agents that any minimal mechanism with limited knowledge must allow. To the best of our knowledge, no such characterization has ever been proposed for the peer prediction setting. Furthermore, we provide a mechanism that reaches the lower bound on the number of dishonest agents. Due to the fact that the bound is logarithmic in the number of reports, aggregated reports converge to the true aggregate. Unlike the mechanism of (Jurca and Faltings 2008; Faltings, Jurca, and Radanovic 2017), that also has a goal of eliciting an accurate aggregate, our mechanism does not require agents to learn from each other’s reports.

The full proofs to our claims can be found in the appendix.

Formal Setting

We study a standard peer prediction setting where agents are assumed to have a *common belief* regarding their private information (Miller, Resnick, and Zeckhauser 2005; Prelec 2004). In the considered setting, a mechanism has almost no knowledge about the agents’ belief structure, which makes the elicibility of truthful information more challenging (e.g., (Radanovic and Faltings 2013)). We define our setting as follows.

There are $n \gg 1$ agents whose arrival to the system is stochastic. We group agents by their arrival so that each group has a fixed number of $k \ll n$ agents, and we consider participation period of a group as a time t .

To describe how agents form beliefs about their private information, we introduce a state ω , which is a random variable that takes values in set Ω , which is assumed to be a (real) interval. We denote the associated distribution by $p(\omega)$, and assume that $p(\omega) > 0$ for all $\omega \in \Omega$.

An agent’s private information, here called *signal*, is modeled with a generic random variable X that takes values in a finite discrete set $\{0, 1, \dots, m - 1\}$ whose generic values are denoted by x, y, z , etc. For each agent i , her signal is generated independently according to a distribution $\Pr(X_i|\omega)$ that depends on state variable ω . This distribution is common for agents, i.e., $\Pr(X_i|\omega) = \Pr(X_j|\omega)$ for two agents

i and j , and it is *fully mixed*, i.e., for all $x \in \{0, \dots, m - 1\}$ it holds that $\Pr(X = x|\omega) > 0$. Furthermore, we assume that private signals are *stochastically relevant* (Miller, Resnick, and Zeckhauser 2005), meaning that posterior distributions $\Pr(X_j|X_i = x)$ and $\Pr(X_j|X_i = y)$ (obtained from $\Pr(X_i|\omega)$, $\Pr(X_j|\omega)$ and $p(\omega)$) differ for at least one value of X_j whenever $x \neq y$. Agents share a common belief about the model parameters of the model ($\Pr(X|\omega)$ and $p(\omega)$), so we denote these beliefs in the same way.

Agents report their private information (signals) to a mechanism, for which they get compensations in terms of rewards. Agents might not be honest, so to distinguish the true signal X from the reported one, we denote reported values by Y . Since our main result depends on agents’ coordination, we also introduce a noise parameter that models potential imperfections in reporting strategies. In particular, we assume that with probability $1 - \epsilon \in (0, 1]$ an agent is *rational* and reports a value that maximizes her expected payoff, while otherwise she heuristically reports a random value from $\{0, \dots, m - 1\}$. Notice that we do not consider adversarial agents. Furthermore, while in the development of our formal results we assume that ϵ is not dependent on X_i , we also show how to apply our main mechanism when such a bias exists (Section *Additional Considerations*).

A mechanism needs not know the true value of ϵ ; it only needs to have an estimate $\hat{\epsilon}$ that is in expectation equal to ϵ , and we show how to obtain $\hat{\epsilon}$ from the reports. Furthermore, the belief of a rational agent i incorporates the fact that a peer report Y_j is noisy, which means that $\Pr(Y_j = x|X_i) = (1 - \epsilon) \cdot \Pr(\bar{X}_j = x|X_i) + \frac{\epsilon}{m}$, where \bar{X}_j is the value that a rational peer would report.

Beliefs about an agent’s signal or her report are defined on the probability simplex in m -dimensional space, that we denote by \mathcal{P} . To simplify the notation for beliefs, we often omit X and Y symbols. In particular, instead of using $\Pr(X = x|\omega)$, we simply write $\Pr(x|\omega)$, or instead of using $\Pr(X_j = y|X_i = x)$, we write $\Pr(y|x)$.

The payments of a mechanism are denoted by τ and they are applied on each group of k agents separately. We are interested in 1-peer payment mechanisms that reward an agent i by using her one peer j , i.e., the reward function is of the form $\tau(Y_i, Y_j)$. As shown in (Radanovic and Faltings 2013), this restriction does not limit the space of strictly incentive compatible mechanisms when agents’ beliefs are not narrowed by particular belief updating conditions. Furthermore, we distinguish the notion of a mechanism, here denoted by \mathcal{M} , from a peer prediction payment function τ because different payment functions could be used on different groups of agents, i.e., at different time periods t .

Solution concept. From the perspective of rational agents, our setting has a form of a Bayesian game, hence we explore strategy profiles that are Bayesian-Nash equilibria. We are particularly interested in strict equilibria, in which rational agents have strict incentives not to deviate. Any mechanism that adopts honest reporting as a strict Bayesian-Nash equilibrium is called strictly Bayesian-Nash *incentive compatible* (BNIC).

Our Approach

Let us begin by describing our approach in dealing with the impossibility of truthful minimal knowledge-bounded elicitation. A mechanism that we are building upon is described by the payment rule:

$$\tau(x, y, P) = d + c \cdot \begin{cases} \frac{1}{P(x)} & \text{if } x = y \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

and is called *the peer truth serum* (PTS) (Faltings et al. 2014). P is a fully mixed distribution that satisfies:

$$\frac{\Pr(x|x)}{P(x)} > \frac{\Pr(y|x)}{P(y)}, y \neq x \quad (2)$$

Provided that other rational agents are honest, the expected payoff of a rational agent with signal x for reporting y is:

$$c \cdot \frac{(1 - \epsilon) \cdot \Pr(y|x) + \frac{\epsilon}{m}}{P(y)} + d$$

Selecting proper values for c and d is an orthogonal problem to the one addressed in the paper, and is typically achieved using a separate mechanism (Radanovic, Faltings, and Jurca 2016), a pre-screening process (Dasgupta and Ghosh 2013), or by learning (Liu and Chen 2016). However, we do set $d = -c \cdot \frac{\epsilon}{m \cdot P(y)}$ so that the expected payoff is proportional to $\frac{\Pr(y|x)}{P(y)}$ (because $\hat{\epsilon} = \mathbb{E}(\epsilon)$). In other words, we remove an undesirable skew in agents' expected payoffs that might occur due to the presence of non-strategic reports.³

Without additional restrictions on agents' beliefs, it is possible to show that the PTS mechanism is uniquely truthful (Frongillo and Witkowski 2016). Condition (2) is called the *self-predicting condition*, and it is crucial for ensuring the truthfulness of PTS.⁴ We say that a distribution P is *informative* if it satisfies the self-predicting condition. Instead of assuming that a specific *a priori* known P satisfies condition (2), we show that there always exists a certain set of distribution functions for which the condition holds, and although this set is initially not known, we show that one can learn it by examining the statistics of reported values for different reporting strategies.

Phase Transition Diagram

We illustrate the reasoning behind our approach and a novel mechanism on a binary answer space $\{0, 1\}$. In this case, it has been shown that if we set $P(x)$ to an agent's prior belief $\Pr(x)$, the self-predicting condition is satisfied, and, consequently, the PTS mechanism is BNIC (Witkowski 2014). However, in our setting, a mechanism has no knowledge about $\Pr(x)$.

Consider what happens when $P(x)$ is much smaller than $\Pr(x)$. For signal value $y \neq x$, this means that $P(y)$ is much

³Furthermore, notice that by setting c proportional to $\prod_x P(x)$, we can bound PTS payments so that they take values in $[0, 1]$.

⁴The condition is typically defined for $P(x)$ equal to the prior $\Pr(x)$ (Jurca and Faltings 2011), but we generalize it here.

larger than $\Pr(y)$. If an agent observes x , her expected payoff when everyone is truthful is proportional to:

$$\frac{\Pr(x|x)}{P(x)} > \frac{\Pr(x|x)}{\Pr(x)} > \frac{\Pr(y|x)}{\Pr(y)}$$

where the last inequality is due to the self-predicting condition. Therefore, agents who observe x are incentivized to report it. However, agents who observe y might not be incentivized to report truthfully, because if $\frac{\Pr(x)}{P(x)} > \frac{\frac{\Pr(y|y)}{\Pr(y)}}{\frac{\Pr(x|y)}{\Pr(x)}}$, we have:

$$\frac{\Pr(x|y)}{P(x)} = \frac{\Pr(x|y)}{\Pr(x)} \cdot \frac{\Pr(x)}{P(x)} > \frac{\Pr(y|y)}{P(y)}$$

In this case, one would naturally expect that both observations x and y lead to report x , and it is easy to verify that this is an equilibrium of the PTS mechanism. Namely, the expected payoffs for reporting x only increase if more agents report x . Similarly, when $P(x)$ is much larger than $\Pr(x)$, one would expect that agents would report y .

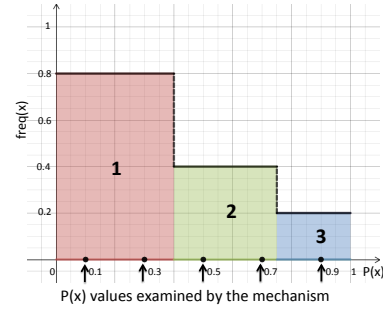


Figure 1: Frequency of reports equal to x for different $P(x)$.

With this reasoning, we can construct a *phase transition* diagram, that shows how the expected frequency of reports equal to x changes with the increase of $P(x)$, for a fixed posterior beliefs $\Pr(\cdot|x)$ and $\Pr(\cdot|y)$, $y \neq x$. The diagram is shown in Figure 1, and it has three phases:

- Phase 1, in which agents are not truthful and report x .
- Phase 2, in which agents are truthful.
- Phase 3, in which agents are not truthful and report y .

Notice that not all reports are equal to x in phase 1 nor equal to y in phase 3. This is due to the presence of noisy reports. However, noisy reports are unbiased, so the frequency $freq_2(x)$ of the truthful reporting phase is by Euclidian distance closer to $\frac{1}{2}$ than are the frequencies of the other two phases:

$$\left(\frac{1}{2} - freq_2(x)\right)^2 < \left(\frac{1}{2} - freq_i(x)\right)^2$$

where $i \in \{1, 3\}$, which gives us:

$$freq_2(x) \cdot (1 - freq_2(x)) > freq(x)_i \cdot (1 - freq_i(x))$$

As the expression also holds for signal $y \neq x$, it follows that the disagreement among reports, i.e., probability that the two reports do not match, is (strictly) maximized in the truthful

reporting phase. Therefore, we can use the disagreement as an indicator of whether agents are truthful or not.

Furthermore, if $\hat{\epsilon}$ needs to be obtained from the reports, it is enough to acquire responses of k agents using PTS that has $d = 0$ and P such that agents are clearly incentivized to report a specific value. For example, one can set $P(x)$ to a small value, and define $\hat{\epsilon} = 1 - \text{num}(x)/k$, where $\text{num}(x)$ is the number of agents who reported x . Notice that $\hat{\epsilon}$ is in expectation equal to $1 - \text{freq}_1(x) = \epsilon$, and this generalizes to the non-binary case (but also ϵ that depends on x (biased noise)).

Mechanism: Adaptive PTS (AdaPTS)

Based on the previous observations, we now construct a novel elicitation mechanism: AdaPTS. The first step of the AdaPTS mechanism is to divide probability simplex \mathcal{P} into regions and accordingly sample from each region one *fully mixed* representative P . In Figure 1, these representatives are shown as black points on the horizontal axis and the division of the simplex is done uniformly. The granularity of the division should be fine enough so that at least one representative P falls into the truthfulness phase. To achieve this, the PTS mechanism must have some knowledge about the agents' belief structure, but this knowledge can be very limited. For example, to properly divide \mathcal{P} , it is enough to know the lower bound on the size of the region that contains distributions P for which PTS is BNIC.

Furthermore, in such a discretization, one can always choose representative P distributions that are in the interior of the probability simplex, thus avoiding potential divisions by 0 in equation (1). Notice that it is also possible to bound payments to a desired interval by choosing an appropriate value of c (see Footnote 2).

Now, the AdaPTS mechanism should define payment function τ before each time step t , i.e., before a considered group of k agents start submitting their reports. We want to maximize the number of honest agents, without knowing for which representative distributions P agents are honest. This can be translated to a (stochastic) multi-armed bandit (MAB) setting⁵ (e.g., see (Agrawal 1995; Auer, Cesa-Bianchi, and Fischer 2002; Audibert, Munos, and Szepesvári 2009; Audibert and Bubeck 2010; Garivier and Cappé 2011)) with arms defined as representative distributions P and the optimization goal defined as maximizing the number of honest agents.

As argued in the previous paragraph, the latter is the same as maximizing the disagreement among reports. More precisely, we define our objective function (feedback of MAB) as an indicator function that counts the disagreements among the reports of k agents:

$$I(Y_1, \dots, Y_k) = \frac{1}{k \cdot (k-1)} \sum_i \sum_{j \neq i} \mathbb{1}_{Y_i \neq Y_j} \quad (3)$$

⁵A most basic K -armed bandit problem is defined by random variables $r_{i,t}$, where i represent the arm of a bandit (gambling machine) and $r_{i,t}$ represents the reward (feedback) obtained by pulling the arm i at time step t . The goal is to maximize the total reward by sampling one arm at each time step.

Notice that the indicator function depends on a chosen representative distribution P , while its associated distribution is dependent on agents' strategies and the underlying distribution from which agents' private signals are sampled. Therefore, at time step t , AdaPTS considers a group of k agents, selects a representative P distribution according to a MAB algorithm, and scores agents using the PTS mechanism with the chosen representative P . After receiving the reports of agents, the mechanism updates the parameters of the MAB algorithm.

Although we could use any MAB algorithm with desirable regret features, in the following text we restrict our attention to UCB1 (Auer, Cesa-Bianchi, and Fischer 2002). Algorithm 1 depicts the pseudocode of AdaPTS based on the UCB1 algorithm (Auer, Cesa-Bianchi, and Fischer 2002). Function *Discretize* returns the set of representative P distributions for a given granularity γ , e.g., by uniformly discretizing the probability simplex as shown in Figure 1. Function *Reports* collects the reports of k agents whose rewards are then calculated using PTS mechanism with parameter P , where the peer of agent i is determined by index $(i+1) \bmod k$. Function *EpsilonEstimate* estimates the value ϵ , e.g., by acquiring responses of k agents for extremal values of P , as described in the previous subsection.

Algorithm 1: AdaPTS

Data: Granularity $\gamma > 0$, Scaling c ;

begin

$RP = \text{Discretize}(\mathcal{P}, \gamma)$;

$\hat{\epsilon} = \text{EpsilonEstimate}(RP, k)$;

for $t = 1$ **to** $t = |RP|$ **do**

$P = RP[t]$;

$\mathbf{x}[1 : k] = \text{Reports}(P, k)$;

$\text{reward}[i \in [1 : k]] =$

$c \cdot \frac{\mathbb{1}_{\mathbf{x}[i] = \mathbf{x}[(i+1) \bmod k]}}{P(\mathbf{x}[i])} - c \cdot \frac{\hat{\epsilon}}{m \cdot P(\mathbf{x}[i])}$;

$I[P] = \frac{1}{k \cdot (k-1)} \sum_{i, j \neq i \in [1:k]} \mathbb{1}_{\mathbf{x}[i] \neq \mathbf{x}[j]}$;

$N[P] = 1$;

for $t = |RP| + 1$ **to** $t = T$ **do**

$P =$

$RP \left[\arg \max_i \left\{ \frac{I[RP[i]]}{N[RP[i]]} + \sqrt{\frac{2 \cdot \ln(t-1)}{N[RP[i]]}} \right\} \right]$;

$\mathbf{x}[1 : k] = \text{Reports}(P, k)$;

$\text{reward}[i \in [1 : k]] =$

$c \cdot \frac{\mathbb{1}_{\mathbf{x}[i] = \mathbf{x}[(i+1) \bmod k]}}{P(\mathbf{x}[i])} - c \cdot \frac{\hat{\epsilon}}{m \cdot P(\mathbf{x}[i])}$;

$I[P] = I[P] + \frac{1}{k \cdot (k-1)} \sum_{i, j \neq i \in [1:k]} \mathbb{1}_{\mathbf{x}[i] \neq \mathbf{x}[j]}$;

$N[P] = N[P] + 1$;

Analysis

We first start by examining particular properties of our setting, which imply the difficulty of our problem and also lead us towards our main results. The major technical difficulty is to show that there exists a distribution P for which mechanism (1) is truthful. This is not a trivial statement, since the original PTS mechanism requires an additional condition to

hold, which is not necessarily satisfied in our setting. Furthermore, we also need to show that (3) is an appropriate indicator function. Given these two results, we can apply the results from multi-armed bandit literature ((Lai and Robbins 1985; Auer, Cesa-Bianchi, and Fischer 2002)) to derive the logarithmic bounds on the dishonesty limit.

Correlation Among Signal Values

The first property we show is that there exist a limit on how different signal values can be correlated in terms of agents posterior beliefs. In particular, if an agent endorses information x , there is an upper bound on the value of her belief about a peer agent endorsing information $y \neq x$.

Lemma 1. *We have:*

$$\Pr(x|x) > \Pr(x), \forall x$$

Furthermore, it holds that:

$$\Pr(x|x) \cdot \Pr(y|y) > \Pr(y|x) \cdot \Pr(x|y), \forall y \neq x$$

or more generally:

$$\Pr(x_1|x_1) \cdots \Pr(x_{m'}|x_{m'}) > \Pr(x_2|x_1) \cdots \Pr(x_1|x_{m'})$$

where $1 < m' \leq m$, $x_i \in \{0, \dots, m-1\}$, $x_i \neq x_j$ for $i \neq j$.

Mechanisms with Limited Knowledge

The second property is that the truthful elicitation of all private signals is not possible if a mechanism has no knowledge about agents' belief structure. This follows from Theorem 1 presented in (Radanovic and Faltings 2013), which states that it is not possible to design a minimal payment scheme that truthfully elicits private signals of all agents. While the result was obtained for a slightly different information elicitation scenario, where no particular belief model is assumed, it is easy to verify that it holds in our setting as well (the proof does not use anything contradictory to our setting). We explicitly state the impossibility of truthful information elicitation due to its importance for the further analysis.

Theorem 1. (Radanovic and Faltings 2013) *There exists no payment function τ that is BNIC for every belief model that complies with the setting.*

Even if a mechanism does have some information about agents, the result of Theorem 1 is likely to hold if this knowledge is limited. We, therefore, define knowledge-bounded mechanisms as mechanisms whose information about agents is not enough to construct a BNIC payments function for all admissible belief models.

Definition 1. *An information structure is a limited knowledge if one cannot construct a payment function τ that is BNIC for every belief model that complies with the setting.*

The AdaPTS mechanism assumes that a given granularity structure of probability simplex \mathcal{P} contains a representative distribution P for which the PTS payment rule is BNIC. We show in the following subsections that one can always partition probability simplex \mathcal{P} to obtain a desirable granularity structure. Moreover, the following lemma shows that the granularity structure of AdaPTS is not in general sufficient to construct a BNIC payment rule.

Lemma 2. *The information structure that AdaPTS has about agents' beliefs is allowed to be a limited knowledge.*

Dishonesty Limit

While one cannot achieve incentive compatibility using peer prediction with limited knowledge, dishonest responses could be potentially useful for a mechanism to learn something about agents. This was noted in (Jurca and Faltings 2008), where the mechanism outputs a publicly available statistic that converges to the desirable outcome - true distribution of private signals. The drawback of the mechanism is that it relies on agents being capable of learning from each other's responses, meaning that they update their beliefs by analyzing the changes in the public statistic.

Our approach is different. By inspecting agents' responses, we aim to learn which incentives are suitable to make agents respond truthfully. To quantify what can be done with such an approach, we define *dishonesty limit*.

Definition 2. *Dishonesty limit (DL) is the minimal expected number of dishonest reports in any Bayesian-Nash equilibrium of any minimal mechanism with a limited knowledge. More precisely:*

$$DL = \min_{\mathcal{M}} \min_{\sigma} d(\mathcal{M}, n, \sigma)$$

where $d(\mathcal{M}, n, \sigma)$ is the expected number of dishonest agents in an equilibrium strategy profile σ of mechanism \mathcal{M} with n agents.

The following lemma establishes the order of the lower bound of DL. Its proof is based on the result of (Lai and Robbins 1985), while the tightness of the result is shown in Section *Main Results*.

Lemma 3. *The dishonesty limit is lower bounded by:*

$$DL \geq \Omega(\log(n))$$

Proof. By Definition 1, there exists no single mechanism that incentivizes agents to report honestly if their belief model is arbitrary. Suppose now that we have two mechanisms τ_1 and τ_2 that are BNIC under two different (complementary) belief models, so that a particular group of agents is truthful only for one mechanism. We can consider this situation from the perspective of a meta-mechanism \mathcal{M} that has to choose between τ_1 and τ_2 . At a time-step t , \mathcal{M} obtains k reports — feedback, which, in general, is insufficient for determining whether agents lied or not because agents' observations are stochastic, while their reports contain noise. Therefore, the problem of choosing between τ_1 and τ_2 is an instantiation of a multi-armed bandit problem (see Section *Mechanism: Adaptive PTS (AdaPTS)*). Since, in general, any MAB algorithm pulls suboptimal arms $\Omega(\log(N))$ number of times in expectation where N is the total number of pulls (e.g., see (Lai and Robbins 1985)), we know that meta mechanism \mathcal{M} will in expectation choose the wrong (untruthful) payments at least $O(\log n)$ times. This produces $O(\log n)$ untruthful reports in expectation because non-truthful payments are not truthful for at least one signal value and each signal value has strictly positive probability of being endorsed by an agent. \square

Existence of an Informative Distribution

The PTS mechanism is BNIC if the associated distribution P satisfies the self-predicting condition, i.e., if P is informative. We now turn to the crucial property of our setting:

Proposition 1. *There exists a region $\mathcal{R} \subset \mathcal{P}$ in probability simplex \mathcal{P} , such that PTS is BNIC for any $P \in \mathcal{R}$.*

Indicator Function

Now, let us define a set of lying strategies that have the same structure and in which agents report only a strict subset of all possible reported values. For example, if possible values are $\{0, 1, 2, 3\}$ and agents are not incentivized to report honestly value 3, then a possible lying strategy could be to report honestly values 0, 1, 2, and instead of honestly reporting 3, agents could report 2.

Definition 3. *Consider a non-surjective function $\rho : \{0, \dots, m-1\} \rightarrow \{0, \dots, m-1\}$. A reporting strategy profile is non-surjective if a report of a rational agent with private signal X is $Y = \rho(X)$.*

Non-surjective strategies also include those that most naturally follow from a simple best response reasoning: agents who are not incentivized to report honestly, deviate by misreporting, which necessarily reduces the set of values that rational agents report. This type of agents' reasoning basically corresponds to the *player inference* process explained in (Waggoner and Chen 2014). Without specifying how agents form their reporting strategy, we show that in PTS there always exists an equilibrium non-surjective strategy profile. In Subsection *Allowing Smoother Transitions Between Phases*, we discuss how to use our approach when agents are not perfectly synchronized in adopting non-surjective strategies.

Proposition 2. *For any fully mixed distribution P , there exists a non-surjective strategy profile that is a strict equilibrium of the PTS mechanism.*

Notice that even for non-surjective strategy profiles, the set of reported values that a mechanism receives does not reduce, i.e., it is equal to $\{0, \dots, m-1\}$. This follows from the fact that not all agents are rational in a sense that they comply with a prescribed strategy profile (i.e., some report random values instead). Nevertheless, the statistical nature of reported values change: reports received by the mechanisms have smaller number of disagreements.

Lemma 4. *The expected value $\mathbb{E}(I(Y_1, \dots, Y_k))$ of the indicator function defined by (3) is strictly greater for truthful reporting than for any non-surjective strategy profile.*

Main Results

From Proposition 1, Proposition 2, and Lemma 4, we obtain the main property of the AdaPTS mechanism: its ability to substantially bound the number of dishonest reports in the system.

Theorem 2. *There exists a strict equilibrium strategy profile of the AdaPTS mechanism such that the expected number of non-truthful reports is of the order of $O(\log n)$.*

Proof. Consider a reporting strategy in which agents are honest whenever P is such that truthful reporting is a strict Bayesian-Nash equilibrium of PTS (by Proposition 1, such P always exists), and otherwise they use an equilibrium non-surjective strategy profile (which by Proposition 2 always

exists). We use the result that the UCB1 algorithm is expected to pull a suboptimal arm $\log(N)$ times, where N is the total number of pulls (Auer, Cesa-Bianchi, and Fischer 2002). By Lemma 4, the representative P of a truthful reporting region is an optimal arm, while the representative P of a non-truthful region is a suboptimal arm. Furthermore, the number of pulls in our case corresponds to n/k , where n is the total number of agents and k is the number of agents at time period t . Since $k \ll n$ is a fixed parameter, the expected number of lying agents is of the order of $O(\log(n))$. \square

Notice that we have not specified an exact equilibrium strategy that satisfies the bound of the theorem. For the theorem to hold, it suffices that agents adopt truthful reporting when P in AdaPTS is such that PTS is BNIC, while they adopt any non-surjective strategy profile when P is such that PTS is not BNIC. As explained in the previous section, a simple best response reasoning can lead to such an outcome.

Since AdaPTS is allowed to have a bounded knowledge information structure (Lemma 2), from Theorem 2 it follows that the dishonesty limit is upper bounded by $O(\log n)$. From Lemma 3, we know that the dishonesty limit is lower bounded by $\Omega(\log n)$. Therefore:

Theorem 3. *The dishonesty limit is:*

$$DL = \Theta(\log n)$$

Importance of the results. With the dishonesty limit concept, we are able to quantify what is possible in the context of minimal elicitation with limited knowledge. An example of an objective that is possible to reach with partially truthful mechanisms is elicitation of accurate aggregates. In particular, suppose that the goal is to elicit a distribution of signal values. From Theorem 2, we know that this is achievable with AdaPTS. Namely, if we denote the normalized histogram of reports by $H_{reports}$ and the normalized histogram of signal values by H_{values} , then it follows from the theorem that their expected difference is bounded by:

$$\mathbb{E}\left(\sum_x |H_{reports}(x) - H_{values}(x)|\right) \leq \frac{O(\log n)}{n}$$

which approaches 0 as n increases. Therefore, although the truthfulness of all agents is not guaranteed, the aggregate obtain by the partially truthful elicitation converges to the one that would be obtained if all agents were honest.

Additional Considerations

In this section, we discuss how to make our mechanism AdaPTS applicable to the situations when there exist biases in reporting errors (i.e., ϵ parameter is biased towards a particular value) or the phase transitions are smoother (e.g., because reporting strategies are not perfectly synchronized).

Allowing Bias in the Non-strategic Reports. Allowing biases in reporting errors is important for cases when some agents do not strategize w.r.t. parameter P , e.g., these agents are truthful regardless of the payment function, or they report heuristically without observing their private signal (for example, a fraction of agents reports 0, whereas the other

agents are strategic). An example of the phase transition diagram that incorporates a bias is shown in Figure 2. Since the disagreement is highest for phase 1 of the diagram, function (3) is not a good indicator of agents' truthfulness. However, frequency $freq_2(x)$ of the truthful phase is the closest one to the average of $f^m(x) = freq_1(x)$ and $f_m(x) = freq_3(x)$. By the same reasoning as in Section *Our Approach*, the expression $(f^m(x) - freq_i(x)) \cdot (freq_i(x) - f_m(x))$ is maximized for $i = 2$. This leads us the following indicator function:

$$I(Y_1, \dots, Y_k) = \sum_x (f^m(x) - \hat{f}(x)) \cdot (\hat{f}(x) - f_m(x))$$

where \hat{f} is the frequency of reports equal to x among $\{Y_1, \dots, Y_k\}$. The problem, however, is that the AdaPTS mechanism does not know f^m and f_m . Nevertheless, it can estimate them in an online manner from agents' reports.

To adjust AdaPTS, we first define time intervals $\theta_i = \{2^i, \dots, 2^{i+1}\}$ and to each time interval associate a different UCB1 algorithm. For periods $t \in \theta_i$, we run a (separate) UCB1 algorithm with the indicator function that uses estimators $\hat{f}_m^i(x)$ and $\hat{f}^i(x)$. The estimators are initially set to $\hat{f}_m^i(x) = 0$ and $\hat{f}^i(x) = 1$ for all values x , and they change after each time interval θ_i . More precisely, they are updated by finding respectively the minimum and the maximum frequency of reports equal to x among all possible arms (representative distributions P). Since UCB1 sufficiently explores suboptimal arms, the estimates $\hat{f}^i(x)$ and $\hat{f}_m^i(x)$ become reasonable accurate at some point, which implies a sublinear number of dishonest reports for a longer elicitation period.

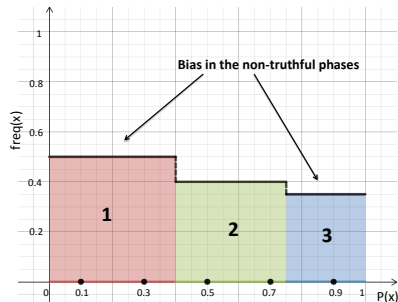


Figure 2: Phase transition diagram with biased phases.

Allowing Smoother Transitions Between Phases.

Agents may not be perfectly synchronized in changing between phases. Nonetheless, we can expect that the resulting behaviour would produce a similar phase transition diagram, as illustrated in Figure 3. However, the simple indicator function defined by expression (3) is no longer a suitable choice for detecting the truthful reporting phase.

In order to see this, we have added additional representative distributions P . Notice that we obtain the same disagreement for $P(x) = 0.4$ as for $P(x) = 0.6$. However, $P(x) = 0.6$ is a better choice, because $P(x) = 0.4$ belongs to a transition phase where a high level of disagreement is due to asynchronous behaviour of agents. Notice that the phase diagram experiences rapid changes in transitions between two phases. This means that we can avoid selection

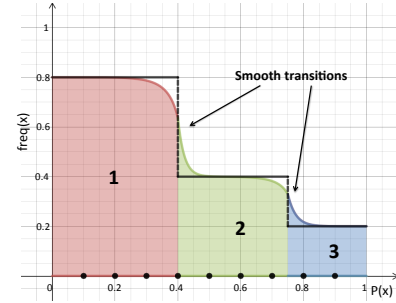


Figure 3: Phase transition diagram with smooth transitions.

of undesirable P distributions by introducing a proper regularization term. That is, we can separate k agents that arrive at time t into two groups, G_1 and G_2 , and reward each group with a slightly different P from the one selected by UCB1. For example, if $P(x) = 0.4$ is selected, we could reward one group using PTS with $P(x) = 0.35$ and the other group using PTS with $P(x) = 0.45$. If group G_1 has agents $\{a_1, \dots, a_4\}$ and group G_2 has agents $\{b_1, \dots, b_4\}$, then a possible indicator function could be:

$$I(Y_1, \dots, Y_k) = \underbrace{\sum_{a_i, a_j \in G_1} \mathbb{1}_{Y_{a_i} \neq Y_{a_j}} + \sum_{b_i, b_j \in G_2} \mathbb{1}_{Y_{b_i} \neq Y_{b_j}}}_{\text{disagreement}} - \underbrace{\lambda \cdot (\mathbb{1}_{Y_{a_1} \neq Y_{a_2}} - \mathbb{1}_{Y_{b_1} \neq Y_{b_2}}) \cdot (\mathbb{1}_{Y_{a_3} \neq Y_{a_4}} - \mathbb{1}_{Y_{b_3} \neq Y_{b_4}})}_{\text{regularization}}$$

where $\lambda > 0$ is the regularization factor. The regularization term is in expectation equal to the square of the difference between the expected disagreement of agents in group G_1 and the expected disagreement of agents in group G_2 . Some insight on how to adjust λ might be a priori needed, but this information is a limited knowledge. With this modification of the indicator function, we can apply AdaPTS.

Conclusion

We investigated the asymptotic behavior of partially truthful minimal peer prediction mechanisms with a limited knowledge. As shown by Theorem 3, any such mechanism results in $O(\log n)$ redundant (non-truthful) reports. In contrast, one of the most known knowledge-bonded elicitation mechanism, Bayesian Truth Serum (Prelec 2004), elicits from each agent her signal value and her prediction about other agents, having in total $O(n)$ additional reports. Thus, our results quantify the necessary overhead when the minimality in reported information and the knowledge of a mechanism is preferred to full truthfulness. One of the most important future steps would be to make the mechanism robust in terms of collusion resistance (e.g., measured using replicator dynamics (Shnayder et al. 2016b)), which, in general, can be challenging even for a more robust settings (Gao, Wright, and Leyton-Brown 2016).

Acknowledgments This work was supported in part by the Swiss National Science Foundation (Early Postdoc Mobility fellowship).

References

- [Agrawal 1995] Agrawal, R. 1995. Sample mean based index policies with $o(\log n)$ regret for the multi-armed bandit problem. *Advances in Applied Probability* 27(4):1054–1078.
- [Audibert and Bubeck 2010] Audibert, J.-Y., and Bubeck, S. 2010. Regret bounds and minimax policies under partial monitoring. *J. Mach. Learn. Res.* 11:2785–2836.
- [Audibert, Munos, and Szepesvári 2009] Audibert, J.-Y.; Munos, R.; and Szepesvári, C. 2009. Exploration-exploitation tradeoff using variance estimates in multi-armed bandits. *Theor. Comput. Sci.* 410(19):1876–1902.
- [Auer, Cesa-Bianchi, and Fischer 2002] Auer, P.; Cesa-Bianchi, N.; and Fischer, P. 2002. Finite-time analysis of the multiarmed bandit problem. *Machine Learning* 47(2-3):235–256.
- [Dasgupta and Ghosh 2013] Dasgupta, A., and Ghosh, A. 2013. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd ACM International World Wide Web Conference (WWW'13)*.
- [Faltings and Radanovic 2017] Faltings, B., and Radanovic, G. 2017. *Game Theory for Data Science: Eliciting Truthful Information*. Morgan & Claypool Publishers.
- [Faltings et al. 2014] Faltings, B.; Pu, P.; Tran, B. D.; and Jurca, R. 2014. Incentives to counter bias in human computation. In *Proceedings of the Second AAAI Conference on Human Computation and Crowdsourcing (HCOMP'14)*.
- [Faltings, Jurca, and Radanovic 2017] Faltings, B.; Jurca, R.; and Radanovic, G. 2017. Peer truth serum: Incentives for crowdsourcing measurements and opinions. *CoRR* abs/1704.05269.
- [Frongillo and Witkowski 2016] Frongillo, R., and Witkowski, J. 2016. A geometric method to construct minimal peer prediction mechanisms. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI'16)*.
- [Gao, Wright, and Leyton-Brown 2016] Gao, A.; Wright, J. R.; and Leyton-Brown, K. 2016. Incentivizing evaluation via limited access to ground truth: Peer-prediction makes things worse. *CoRR* abs/1606.07042.
- [Garcin and Faltings 2014] Garcin, F., and Faltings, B. 2014. Swissnoise: Online polls with game-theoretic incentives. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence (AAAI'14)*.
- [Garivier and Cappé 2011] Garivier, A., and Cappé, O. 2011. The KL-UCB algorithm for bounded stochastic bandits and beyond. In *The 24th Annual Conference on Learning Theory (COLT'11)*, 359–376.
- [Jurca and Faltings 2008] Jurca, R., and Faltings, B. 2008. Incentives for expressing opinions in online polls. In *Proceedings of the 9th ACM conference on Electronic commerce (EC'08)*.
- [Jurca and Faltings 2011] Jurca, R., and Faltings, B. 2011. Incentives for answering hypothetical questions. In *Workshop on Social Computing and User Generated Content*.
- [Kamble et al. 2015] Kamble, V.; Shah, N.; Marn, D.; Parekh, A.; and Ramachandran, K. 2015. Truth serums for massively crowdsourced evaluation tasks. In *the 5th Workshop on Social Computing and User Generated Content*.
- [Kong and Schoenebeck 2016] Kong, Y., and Schoenebeck, G. 2016. Equilibrium selection in information elicitation without verification via information monotonicity. *CoRR* abs/1603.07751.
- [Lai and Robbins 1985] Lai, T. L., and Robbins, H. 1985. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics* 6(1):4–22.
- [Liu and Chen 2016] Liu, Y., and Chen, Y. 2016. Learning to incentivize: Eliciting effort via output agreement. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*.
- [Miller, Resnick, and Zeckhauser 2005] Miller, N.; Resnick, P.; and Zeckhauser, R. 2005. Eliciting informative feedback: The peer-prediction method. *Management Science* 51:1359–1373.
- [Prelec 2004] Prelec, D. 2004. A bayesian truth serum for subjective data. *Science* 34(5695):462–466.
- [Radanovic and Faltings 2013] Radanovic, G., and Faltings, B. 2013. A robust bayesian truth serum for non-binary signals. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI'13)*.
- [Radanovic and Faltings 2014] Radanovic, G., and Faltings, B. 2014. Incentives for truthful information elicitation of continuous signals. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI'14)*.
- [Radanovic, Faltings, and Jurca 2016] Radanovic, G.; Faltings, B.; and Jurca, R. 2016. Incentives for effort in crowdsourcing using the peer truth serum. *ACM Transactions on Intelligent Systems and Technology (TIST)* 7:48:1–48:28.
- [Shnayder et al. 2016a] Shnayder, V.; Agarwal, A.; Frongillo, R.; and Parkes, D. C. 2016a. Informed truthfulness in multi-task peer prediction. In *Proceedings of the 2016 ACM Conference on Economics and Computation (EC'16)*.
- [Shnayder et al. 2016b] Shnayder, V.; Agarwal, A.; Frongillo, R.; and Parkes, D. C. 2016b. Measuring performance of peer prediction mechanisms using replicator dynamics. In *In Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*.
- [Waggoner and Chen 2014] Waggoner, B., and Chen, Y. 2014. Output agreement mechanisms and common knowledge. In *Proceedings of the Second AAAI Conference on Human Computation and Crowdsourcing (HCOMP'14)*.
- [Witkowski and Parkes 2012a] Witkowski, J., and Parkes, D. C. 2012a. Peer prediction without a common prior. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC'12)*.
- [Witkowski and Parkes 2012b] Witkowski, J., and Parkes, D. C. 2012b. A robust bayesian truth serum for small populations. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI'12)*.

[Witkowski 2014] Witkowski, J. 2014. *Robust Peer Prediction Mechanisms*. Ph.D. Dissertation, Albert-Ludwigs-Universität Freiburg: Institut für Informatik.

ATTACHMENT: Partial Truthfulness in Peer Prediction Mechanisms with Limited Knowledge

Proof of Lemma 1

Proof. Using the properties of our model (conditional independence of signal values given ω) and Bayes' rule we obtain:

$$\begin{aligned} \Pr(x|x) &= \int_{\omega} \Pr(x|\omega) \cdot p(\omega|x) d\omega = \int_{\omega} \Pr(x|\omega)^2 \cdot \frac{p(\omega)}{\Pr(x)} d\omega \\ &= \frac{\int_{\omega} \Pr(x|\omega)^2 \cdot p(\omega) d\omega}{\int_{\omega} \Pr(x|\omega) \cdot p(\omega) d\omega} \end{aligned}$$

Jensen's inequality tells us that $\int_{\omega} \Pr(x|\omega)^2 \cdot p(\omega) d\omega \geq (\int_{\omega} \Pr(x|\omega) \cdot p(\omega) d\omega)^2$, with strict inequality when $0 < \Pr(x|\omega) < 1$ (notice that $\Pr(x|\omega)$ is not constant due to stochastic relevance). As $\Pr(x|\omega)$ is fully mixed, we have:

$$\begin{aligned} \Pr(x|x) &> \frac{(\int_{\omega} \Pr(x|\omega) \cdot p(\omega) d\omega)^2}{\int_{\omega} \Pr(x|\omega) \cdot p(\omega) d\omega} > \int_{\omega} \Pr(x|\omega) \cdot p(\omega) d\omega \\ &= \Pr(x) \end{aligned}$$

implying the first statement.

From Bayes' rule it follows that $\Pr(x|x) \cdot \Pr(y|y)$ is equal to:

$$\begin{aligned} &\Pr(x|x) \cdot \Pr(y|y) \\ &= \frac{\int_{\omega} \Pr(x|\omega)^2 \cdot p(\omega) d\omega}{\int_{\omega} \Pr(x|\omega) \cdot p(\omega) d\omega} \cdot \frac{\int_{\omega} \Pr(y|\omega)^2 \cdot p(\omega) d\omega}{\int_{\omega} \Pr(y|\omega) \cdot p(\omega) d\omega} \quad (4) \end{aligned}$$

Similarly, $\Pr(y|x) \cdot \Pr(x|y)$ is equal to:

$$\begin{aligned} &\Pr(y|x) \cdot \Pr(x|y) \\ &= \frac{\int_{\omega} \Pr(y|\omega) \cdot \Pr(x|\omega) \cdot p(\omega) d\omega}{\int_{\omega} \Pr(x|\omega) \cdot p(\omega) d\omega} \cdot \frac{\int_{\omega} \Pr(x|\omega) \cdot \Pr(y|\omega) \cdot p(\omega) d\omega}{\int_{\omega} \Pr(y|\omega) \cdot p(\omega) d\omega} \quad (5) \end{aligned}$$

Notice that expressions (4) and (5) have equal denominators, so we only need to compare nominators. Let A and B be two random variables such that $A = \Pr(x|\omega)$ and $B = \Pr(y|\omega)$. We have:

$$\begin{aligned} &\text{sgn}(\Pr(x|x) \cdot \Pr(y|y) - \Pr(x|x) \cdot \Pr(y|y)) \\ &= \text{sgn}(\mathbb{E}_{\omega}(A^2) \cdot \mathbb{E}_{\omega}(B^2) - \mathbb{E}_{\omega}(A \cdot B)^2) \end{aligned}$$

\mathbb{E}_{ω} is expectation over the distribution $p(\omega)$. Using the Cauchy-Schwarz inequality ($\mathbb{E}_{\omega}(A^2) \cdot \mathbb{E}_{\omega}(B^2) \geq |\mathbb{E}_{\omega}(A \cdot B)|^2$), with strict inequality if $A, B \neq 0$ and $A \neq \lambda \cdot B$, and the fact that A and B are positive random variables that differ due to stochastic relevance, we obtain that $\Pr(x|x) \cdot \Pr(y|y) > \Pr(y|x) \cdot \Pr(x|y)$ for $x \neq y$.

The third claim follows analogously. Namely, the nominator of (4) is in more general form equal to

$\mathbb{E}_{\omega}(A_1^2) \cdots \mathbb{E}_{\omega}(A_{m'}^2)$, while the nominator of (5) is in more general form equal to $\mathbb{E}_{\omega}(A_1 \cdot A_2) \cdots \mathbb{E}_{\omega}(A_{m'} \cdot A_1)$, where $A_i = \Pr(x_i|\omega)$. By applying the Cauchy-Schwarz inequality, we obtain the claim. \square

Proof of Lemma 2

Proof. Consider two arbitrary probability distribution functions P_1 and P_2 , such that: $P_1(x) = P_2(y)$, $P_1(y) = P_2(x)$, and $P_1(z) = P_2(z)$ for $z \neq x$ and $z \neq y$. PTS with P_1 is truthful for belief model defined by $Pr_1(z|z) = P_1(z) + (1 - P_1(z)) \cdot \beta$, $Pr(w|z) = P_1(w) - P_1(w) \cdot \beta$, $\forall z, w \in \{0, \dots, m-1\} : w \neq z$, where $\beta \in (0, 1)$. Similarly, we define posteriors Pr_2 based on P_2 . Notice that $Pr_2(w|z) = Pr_1(u(w)|u(z))$, where $u : \{0, \dots, m-1\} \rightarrow \{0, \dots, m-1\}$:

$$u(z) = \begin{cases} y & \text{if } z = x \\ x & \text{if } z = y \\ z & \text{if } z \neq x \wedge z \neq y \end{cases}$$

Now, suppose that $P_1(x) > P_1(y)$, and set up β such that: $Pr_1(x|x) > Pr_1(y|x)$ and $Pr_1(y|y) < Pr_1(x|y)$. By using the same procedure of proving as in Theorem 1 of (Radanovic and Faltings 2013), we obtain that if a mechanism τ is incentive compatible for both Pr_1 and Pr_2 , then:

$$\begin{aligned} &(Pr_1(x|x) - Pr_1(y|x)) \cdot (\Delta(x) - \Delta(y)) > 0 \\ &(Pr_1(y|y) - Pr_1(x|y)) \cdot (\Delta(x) - \Delta(y)) > 0 \end{aligned}$$

where $\Delta(x) = \tau(x, x) - \tau(x, y)$ and $\Delta(y) = \tau(y, x) - \tau(y, y)$. The two inequalities, however, contradict by the choice of Pr_1 .

Therefore, even though a mechanism might know the size of a region in probability simplex \mathcal{P} that contains distributions P for which PTS is BNIC, it might not be able to construct a BNIC payment rule. \square

Proof of Lemma 3

Proof. By Definition 1, there exists no single mechanism that incentivizes agents to report honestly if their belief model is arbitrary. Suppose now that we have two mechanisms τ_1 and τ_2 that are BNIC under two different (complementary) belief models, so that a particular group of agents is truthful only for one mechanism. We can consider this situation from the perspective of a meta-mechanism \mathcal{M} that has to choose between τ_1 and τ_2 . At a time-step t , \mathcal{M} obtains k reports — feedback, which, in general, is insufficient for determining whether agents lied or not because agents' observations are stochastic, while their reports contain noise. Therefore, the problem of choosing between τ_1 and τ_2 is an instantiation of a multi-armed bandit problem (see Section *Mechanism: Adaptive PTS (AdaPTS)*). Since, in general, any MAB algorithm pulls suboptimal arms $\Omega(\log(N))$ number of times in expectation where N is the total number of pulls (e.g., see (Lai and Robbins 1985)), we know that meta mechanism \mathcal{M} will in expectation choose the wrong (untruthful) payments at least $O(\log n)$ times. This produces $O(\log n)$ untruthful reports in expectation because non-truthful payments are not truthful for at least one signal value and each signal value has strictly positive probability of being endorsed by an agent. \square

Proof of Proposition 1

The proposition is the direct consequence of the two following lemmas.

Lemma 5. *If there exists an informative distribution $P \in \mathcal{P}$, where \mathcal{P} is probability simplex for m -ary signal space, then there exists a region $\mathcal{R} \subset \mathcal{P}$ such that all $P' \in \mathcal{R}$ are informative.*

Proof. Since P is informative:

$$\frac{\Pr(x|x)}{P(x)} > \frac{\Pr(y|x)}{P(y)}$$

for all $x \neq y$. The strictness of the inequality implies that there exists $0 < \delta \ll \min_z Pr(z)$ such that for all $x \neq y$ we have:

$$\frac{Pr(x|x)}{P(x) + \delta} > \frac{Pr(y|x)}{P(y) - \delta}$$

In other words, we have that any $P' \in \mathcal{P}$ for which $|P'(x) - P(x)| \leq \delta$, $x \in \{1, \dots, m\}$, satisfies:

$$\frac{Pr(x|x)}{P'(x)} > \frac{Pr(y|x)}{P'(y)}$$

By putting $\mathcal{R} = \{P' | P' \in \mathcal{P}, \forall x : |P'(x) - P(x)| < \delta\}$, we obtain the claim. \square

Lemma 6. *There exists an informative distribution P .*

Proof. We only need to show the existence of P . Consider a specific signal value $a \in \{0, \dots, m-1\}$ and any other signal value $x \in \{0, \dots, m-1\}$. Let us define P as:

$$P(a) = \frac{1}{\alpha}$$

$$P(x) = \frac{1}{\alpha} \cdot \min_{y \in \{0, \dots, m-1\}} \left[\frac{\Pr(x|x)}{\Pr(y|x)} \cdot \frac{\Pr(y|y)}{\Pr(a|y)} \right]$$

where $\frac{1}{\alpha}$ is a normalization factor so that $\sum_{y \in \{0, \dots, m-1\}} P(y) = 1$. Notice that by Lemma 1:

$$\frac{\Pr(x|x)}{\Pr(y|x)} \cdot \frac{\Pr(y|y)}{\Pr(a|y)} > \frac{\Pr(x|a)}{\Pr(a|a)} \quad (6)$$

holds for any $y \in \{0, \dots, m-1\}$. To prove that P is informative, it is sufficient to show that for any signal values $x \neq z \neq a$ we have:

$$\frac{\Pr(a|a)}{P(a)} > \frac{\Pr(x|a)}{P(x)}$$

$$\frac{\Pr(x|x)}{P(x)} \geq \frac{\Pr(a|x)}{P(a)}$$

$$\frac{\Pr(x|x)}{P(x)} > \frac{\Pr(z|x)}{P(z)}$$

Provided that these inequalities hold, the second one can be made strict (while keeping the other two inequalities strict as well) by reducing all $P(x)$, $x \neq a$, by a small enough value, and then re-normalizing P .

The first inequality follows from $\frac{\Pr(a|a)}{P(a)} = \alpha \cdot \Pr(a|a)$ and inequality (6):

$$\frac{\Pr(x|a)}{P(x)} = \alpha \cdot \frac{\Pr(x|a)}{\min_{y \in \{0, \dots, m-1\}} \left[\frac{\Pr(x|x)}{\Pr(y|x)} \cdot \frac{\Pr(y|y)}{\Pr(a|y)} \right]}$$

$$< \alpha \cdot \frac{\Pr(x|a)}{\frac{\Pr(x|x)}{\Pr(a|x)}} = \alpha \cdot \Pr(a|a) = \frac{\Pr(a|a)}{P(a)}$$

We obtain the second inequality by putting $y = a$:

$$\frac{\Pr(x|x)}{P(x)} = \alpha \cdot \frac{\Pr(x|x)}{\min_{y \in \{0, \dots, m-1\}} \left[\frac{\Pr(x|x)}{\Pr(y|x)} \cdot \frac{\Pr(y|y)}{\Pr(a|y)} \right]}$$

$$\geq \alpha \cdot \frac{\Pr(x|x)}{\frac{\Pr(x|x)}{\Pr(a|x)}} = \alpha \cdot \Pr(a|x) = \frac{\Pr(a|x)}{P(a)}$$

For the third inequality, we use the second inequality and inequality $\Pr(x|x) \cdot \Pr(z|z) > \Pr(z|x) \cdot \Pr(x|z)$, which follows from Lemma 1. We have:

$$\frac{\Pr(z|x)}{P(z)} < \frac{\Pr(x|x) \cdot \Pr(z|z)}{\Pr(x|z)} \cdot \frac{1}{P(z)}$$

$$= \frac{\Pr(x|x) \cdot \Pr(z|z)}{\Pr(x|z)} \cdot \frac{\alpha}{\min_{y \in \{0, \dots, m-1\}} \left[\frac{\Pr(z|z)}{\Pr(y|z)} \cdot \frac{\Pr(y|y)}{\Pr(a|y)} \right]}$$

$$\leq \alpha \cdot \frac{\Pr(x|x) \cdot \Pr(z|z)}{\Pr(x|z) \cdot \left[\frac{\Pr(z|z)}{\Pr(a|z)} \cdot \frac{\Pr(x|x)}{\Pr(a|x)} \right]} = \alpha \cdot Pr(a|x)$$

$$= \frac{\Pr(a|x)}{P(a)} \leq \frac{\Pr(x|x)}{P(x)}$$

\square

Proof of Proposition 2

Proof. One equilibrium non-surjective strategy profile is when agents report value x such that $P(x) = \min_z P(z)$. Let us denote by $\tilde{P}r$ an agent i 's belief regarding the report of her peer agent j , i.e., Y_j , for the considered strategy profile. Notice that $\tilde{P}r(y|x) = \frac{\epsilon}{m} \geq 0$ for $y \neq x$ (due to the reporting noise). However, the strategy profile of reporting x is an equilibrium because the expected value of $\hat{\epsilon}$ is equal to ϵ , so in expectation $c \cdot \frac{\tilde{P}r(y|x)}{P(y)} - d = c \cdot \frac{\epsilon}{m \cdot P(y)} - c \cdot \frac{\epsilon}{m \cdot P(y)} = 0$ for $y \neq x$, while $c \cdot \frac{\tilde{P}r(x|x)}{P(x)} - d > 0$. That is, an agent's expected payment is strictly maximized when she reports x . \square

Proof of Lemma 4

Proof. Since a non-surjective reporting strategy is a non-surjective function of observation X_i , we know that $\rho(X_i)$ takes values in a strict subset of all possible signal values. The disagreement function is linear in $\mathbb{1}_{Y_i \neq Y_j}$, so it is sufficient to show that $\frac{1}{1-\epsilon} \cdot (\mathbb{1}_{Y_i \neq Y_j} - \frac{\epsilon}{m})$ is in expectation greater for truthfulness than for non-surjective reporting strategy profile. In expectation, the expression is equivalent to saying whether two reports of rational agents disagree, which is equal to $1 - p_a$, where p_a is the probability of agreement. The

probability of agreement in a non-surjective strategy profile is equal to:

$$p_{a,ns} = \sum_x \Pr(\{z|\rho(z) = \rho(x)\}, x) > \sum_x \Pr(x, x) = p_{a,t}$$

where the last term is the probability of agreement for truthfulness. Notice that the inequality is strict because in a non-surjective strategy profile there exist x and y for which $\rho(x) = \rho(y)$, and thus, $\Pr(\{z|\rho(z) = \rho(x)\}, x) \geq \Pr(\{x, y\}, x) = \Pr(x, x) + \Pr(y, x) > \Pr(x, x)$. The last inequality follows from agents' beliefs being fully mixed. Since $p_{a,t}$ is strictly smaller than $p_{a,ns}$ for any non-surjective strategy profile, we conclude that the disagreement is strictly greater for truthful reporting than for any other non-surjective strategy profile. \square

Proof of Theorem 2

Proof. Consider a reporting strategy in which agents are honest whenever P is such that truthful reporting is a strict Bayesian-Nash equilibrium of PTS (by Proposition 1, such P always exists), and otherwise they use an equilibrium non-surjective strategy profile (which by Proposition 2 always exists). We use the result that the UCB1 algorithm is expected to pull a suboptimal arm $\log(N)$ times, where N is the total number of pulls (Auer, Cesa-Bianchi, and Fischer 2002). By Lemma 4, the representative P of a truthful reporting region is an optimal arm, while the representative P of a non-truthful region is a suboptimal arm. Furthermore, the number of pulls in our case corresponds to n/k , where n is the total number of agents and k is the number of agents at time period t . Since $k \ll n$ is a fixed parameter, the expected number of lying agents is of the order of $O(\log(n))$. \square