The Limits of Multi-task Peer Prediction

SHURAN ZHENG, FANG-YI YU, and YILING CHEN, Harvard University

Recent advances in multi-task peer prediction have greatly expanded our knowledge about the power of multi-task peer prediction mechanisms. Various mechanisms have been proposed in different settings to elicit different types of information. But we still lack understanding about when desirable mechanisms will exist for a multi-task peer prediction problem. In this work, we study the elicitability of multi-task peer prediction problems. We consider a designer who has certain knowledge about the underlying information structure and wants to elicit certain information from a group of participants. Our goal is to infer the possibility of having a desirable mechanism based on the primitives of the problem.

Our contribution is twofold. First, we provide a characterization of the elicitable multi-task peer prediction problems, assuming that the designer only uses *scoring mechanisms*. Scoring mechanisms are the mechanisms that reward participants' reports for different tasks separately. The characterization uses a geometric approach based on the *power diagram* characterization [5, 13] in the single-task setting. For general mechanisms, we also give a necessary condition for a multi-task problem to be elicitable.

Second, we consider the case when the designer aims to elicit some properties that are linear in the participant's posterior about the state of the world. We first show that in some cases, the designer basically can only elicit the posterior itself. We then look into the case when the designer aims to elicit the participants' posteriors. We give a necessary condition for the posterior to be elicitable. This condition implies that the mechanisms proposed by [11] are already the best we can hope for in their setting, in the sense that their mechanisms can solve any problem instance that can possibly be elicitable.

CCS Concepts: • **Theory of computation → Algorithmic game theory and mechanism design**; **Algorithmic mechanism design**.

Additional Key Words and Phrases: information elicitation, peer prediction, multi-task peer prediction, power diagram, scoring mechanisms

ACM Reference Format:
Shuran Zheng, Fang-Yi Yu, and Yiling Chen. 2021. The Limits of Multi-task Peer Prediction. In *Proceedings of the 22nd ACM Conference on Economics and Computation (EC '21), July 18–23, 2021, Budapest, Hungary.* ACM, New York, NY, USA, 28 pages. https://doi.org/10.1145/3465456.3467642

1 INTRODUCTION

Peer prediction refers to a collection of incentive mechanisms [3, 7, 11, 12, 15–20, 22, 24] that have been designed for the challenging setting where truthful information elicitation about some tasks is desired but the designer has no access to the ground truth (i.e. event outcomes) for incentive alignment. This setting is fundamental to many information elicitation applications such as peer grading, surveys, product reviews, and forecasting for long-term events.

Recent advances in peer prediction have progressed from single-task peer prediction [7, 16–19, 24], where an agent's reward on a task is solely determined by how his report on the task relates to the reports made by peer agents on the same task, to multi-task peer prediction [3, 11, 12, 15, 20, 22], where reports made by peer agents on other tasks can also be used in determining

EC '21, July 18–23, 2021, Budapest, Hungary

ACM ISBN 978-1-4503-8554-1/21/07…$15.00
https://doi.org/10.1145/3465456.3467642

the agent's reward on the task. Multi-task peer prediction mechanisms can often achieve stronger incentive guarantees or require fewer assumptions on the underlying information structure than single-task peer prediction mechanisms, thanks to the additional cross-task information. While the quest for better peer prediction mechanisms is bound to continue, we attempt to understand the limits for designing multi-task peer prediction mechanisms in this work: When is it possible to design a desirable multi-task peer prediction mechanism?

To answer this question, the first thing we may need to understand is: what are the factors that will influence the elicitability of a problem? Our first observation is that the existing multi-task mechanisms (as well as single-task mechanisms) all rely on certain knowledge about the underlying information structure or various assumptions about it. For example, Shnayder et al. [22] required the designer to partially know the correlation between the participants' signals; Kong and Schoenebeck [11] and Liu et al. [15] assumed that the participants' signals are mutually independent conditioned on the unknown ground truth. In addition to the various assumptions about the underlying distribution, we have also seen mechanisms that utilize the structure of the reported information. For example, Kong and Schoenebeck [11] proposed a mechanism that rewards the *point-wise mutual information* between the participants' reports, which can only be computed when the reports are the participants' posteriors about the state.

Our problem becomes more clear. Suppose there is a designer who has certain knowledge about the underlying information structure and she wants to elicit certain information from a group of participants, can we infer the possibility of having a desirable mechanism based on the primitives of the problem before trying to search for mechanisms? The answer to this question may also shed light on the design of new mechanisms. For certain information we want to elicit, what do we have to know about the underlying information structure? Based on our knowledge about the information structure, what kind of information can we possibly elicit?

This problem has been studied in single-task peer prediction if we consider mechanisms that only ask the participants for their signals. Frongillo and Witkowski [5] used a geometric perspective to prove that single-task peer prediction mechanisms that achieve strict truthfulness are equivalent to *power diagrams*. Their result gives a necessary and sufficient condition for a designer with certain knowledge about the participants' posterior beliefs to be able to design a strictly truthful mechanism: the designer should be able to divide a participant' possible posteriors after seeing different signal realizations into different regions, and moreover, these regions must take a particular shape, that of a *power diagram*.

But for multi-task peer prediction, little is known about the exact condition for the existence of strictly truthful mechanisms. As we will show by an example (at the beginning of Section 4), it is possible for the designer to exploit the similarity between the tasks and elicit information that is not elicitable in the single-task setting, assuming that the designer has the same knowledge about the distribution. A natural thought one may have is to view a multi-task problem as a single-task problem in which a participant's report is a combination of reports for multiple tasks. The problem of directly converting a multi-task problem into a single task problem is that a report will have exponentially many possible values and the condition given by Frongillo and Witkowski [5] will involve power diagrams in dimension that grows exponentially in the number of tasks, which may not lead to meaningful results. Even for a constant number of tasks, directly applying their method does not give an easily interpretable characterization.

In this work, we study the elicitability in multi-task peer prediction. For the definition of elicitability, we consider the most basic incentive guarantee that truth-telling is a strict Bayesian Nash equilibrium (BNE). We say that a peer prediction problem is elicitable if there exists a mechanism that guarantees that truth-telling is a strict BNE for any possible underlying information structure. We consider a peer prediction problem to have two primitives. The first one is the designer's

knowledge or assumption about the underlying information structure. The knowledge may have different forms. In this work, we model the designer's knowledge by a set of possible underlying information structures. The designer knows that the underlying information structure must lie in this set, but she does not know which one is the true one. The second primitive is the information that the designer asks each participant to report. In this work, we consider the most general information which can be a function of the information structure and the participant's signal.

Our contribution is twofold. First, we give a characterization of the elicitable multi-task peer prediction problems, assuming that the designer only uses *scoring mechanisms*. Scoring mechanisms are the mechanisms that reward participants' reports for different tasks separately. To our knowledge, all the existing mechanisms that achieve strict truthfulness are scoring mechanisms.[1] We show that a multi-task problem is elicitable if and only if the following two conditions are satisfied: (1) the designer should be able to separate a participant's possible posteriors after seeing different signal realizations using a power diagram, *for any given marginal distribution of other participants' truthful reports*; (2) for different marginal distributions of other participants' truthful reports, the parameters of the power diagrams should be an affine function of the marginal distribution of other participants' truthful reports. For general mechanisms, we give a necessary condition for a multi-task problem to be elicitable. The necessary condition basically says that, given a joint distribution of the participants' reports without naming a participant i's report, the designer should at least be able to label participant i's report based on the distribution.

Second, we consider the case when the designer aims to elicit some properties that are linear in the participant's posterior about the state of the world. We first show that in some cases, the designer basically can only elicit the posterior itself. More specifically, we apply our characterization to the case when there are two participants with signals independent conditioning on the state. If the designer only uses scoring mechanisms, then the only elicitable linear properties of the posterior are the ones that are equivalent to the posterior, assuming that the designer is uncertain about the underlying distribution. We then look into the case when the designer elicits the participants' posteriors. We give a necessary condition for the posterior to be elicitable. This condition implies that the mechanisms proposed by [11] are already the best we can hope for in their setting, in the sense that their mechanisms can work for any problem instance that can possibly be elicitable.

1.1 Related work

The elicitability of peer prediction problems has not received a lot of attention. For the single-task peer prediction, Frongillo and Witkowski [5] use a geometric approach to study necessary and sufficient conditions for the existence of strictly truthful peer prediction mechanisms. However, their characterizations are for single-task mechanisms that only collect agents' signals, but many single-task mechanisms elicit information other than or in addition to agents' signals or require relatively strong assumptions on the underlying information structure. Zhang and Chen [25] also consider the existence of strictly truthful mechanisms. They show stochastic relevance is a necessary condition even for general mechanisms.

Although the problem of elicitability has not been extensively investigated in the peer prediction literature. There is a vast literature on *property elicitation*. We are not able to review all of the works in this area but point the readers to [6] and the references therein. In property elicitation, the designer asks an agent to report a property of a probability distribution. The designer is able to observe a sample drawn from the distribution and then decide the payment based on the report and

[1]The only mechanism that we know does not belong to scoring mechanisms is the Determinant-based Mutual Information mechanism proposed by Kong [8]. The mechanism is not strictly truthful because it cannot distinguish permutation strategies from truth-telling.

the sample. To a certain extent, a peer prediction problem can be viewed as a property elicitation problem in which the other participants' reports is a sample. But the problem is that it may not always be possible to represent the information that the designer wants to elicit as a property of the distribution of other participants' reports.

Finally, we review the existing literature on peer prediction, in both multi-task setting and single-task setting. Also see [4] for a survey of additional results.

Multi-task setting. The multi-task peer prediction problem was first independently introduced and studied by Dasgupta and Ghosh [3] and Witkowski and Parkes [23]. Agents are assigned a batch of a priori similar tasks which require each agents' private information to be a binary signal. Later works extend the setting to multiple-choice questions and design mechanisms that achieve various truthfulness guarantees (dominant truthful, informed truthful) [3, 8, 12, 20, 22]. But none of them is strictly truthful on general distributions, because agents can always relabel their signals. Liu et al. [15] design an approximated dominant truthful mechanism (also approximated strictly truthful) that uses surrogate loss functions as tools to correct for the mistakes in agents' reports. Kong and Schoenebeck [11] study the related goal for forecast elicitation. All the above mechanisms [3, 11, 12, 15, 20, 22] are scoring mechanisms (Definition 2.2) except the DMI mechanism by Kong [8] and the VMI mechanisms by Kong [9].

Single-task setting. Miller et al. [16] introduced the original peer prediction mechanism, which is the first mechanism that has truth-telling as a strict Bayesian Nash equilibrium and does not need verification. However, their mechanism requires the full knowledge of the common prior. Prelec [17] relaxes the full knowledge assumption and designs the first detail-free peer prediction mechanism—Bayesian truth serum (BTS). BTS requires that all agents' signals are symmetric and conditional independent given a latent state. Several other works study the same single-task setting as BTS and devise mechanisms that work on more general underlying information structures [7, 10, 18, 19, 21, 24].

2 PROBLEM DESCRIPTION

Consider a designer who wants to elicit certain information about the state of the world $\omega \in \Omega$ from a group of participants. There are n participants who receive private signals $s_1, \dots, s_n$ respectively with $s_1 \in \mathcal{S}_1, \dots, s_n \in \mathcal{S}_n$. We use S_i to denote the random variable for participant i's signal. The state of the world and the signals follow an unknown underlying distribution $\mu(\omega, s_1, \dots, s_n)$. The designer does not know the true underlying distribution μ, but she may have some information about the structure of the distribution, which allows her to restrict μ to a set $M \subseteq \Delta(\Omega \times \mathcal{S}_1 \times \cdots \times \mathcal{S}_n)$. We assume that the underlying distribution $\mu(\omega, s_1, \dots, s_n)$ is common knowledge for all the participants. But participant i only observes the realization of his own signal s_i and thus his posterior belief about the state and others' signals will be $\mu(\omega, \mathbf{s}_{-i}|s_i)$, where $\mathbf{s}_{-i}$ denote the signals of the participants other than i.

The designer's goal is to elicit certain information from the participants. Participant i will be asked to report a function of his own signal $\mathbf{r}_i = f_{i,\mu}(s_i)$. In this work, we consider functions that are real vectors $\mathbf{r}_i \in \mathbb{R}^L$. Note that the report function can possibly depend on μ. For example, an extensively studied report function is the prediction of the state $\mathbf{r}_i = \mu(\omega|s_i)$; and the well known Bayesian Truth Serum [17] asks for the prediction of other people's signals $\mathbf{r}_i = \mu(\mathbf{s}_{-i}|s_i)$. Throughout the work, we use $\mathbf{r}_i = f_{i,\mu}(s_i)$ to represent participant i's *truthful* report. We denote by $\mathbf{R}_i = f_{i,\mu}(S_i)$ the random variable for participant i's truthful report and denote by $\mathcal{R}_i$ the range of the report function $f_{i,\mu}(s_i)$.

In multi-task peer prediction, the designer elicits information for $T > 1$ i.i.d. tasks. More specifically, we have

$$(\omega^{(1)}, S_1^{(1)}, \ldots, S_n^{(1)}), \ldots, (\omega^{(T)}, S_1^{(T)}, \ldots, S_n^{(T)}) \overset{i.i.d}{\sim} \mu(\omega, s_1, \ldots, s_n)$$

where $(\omega^{(t)}, S_1^{(t)}, \ldots, S_n^{(t)})$ indicates the state and the signals for task t. The designer elicits the same information from a participant for all T tasks, i.e., $\mathbf{r}_i^{(t)} = f_{i,\mu}(s_i^{(t)})$ for the same function $f_{i,\mu}(\cdot)$ across all the tasks. We denote by $\omega^{(1:T)}$, $\mathbf{s}_i^{(1:T)}$ and $\mathbf{r}_i^{(1:T)}$ the vector of the states, the vector of participant i's signal realizations and the vector of participant i's truthful reports for all T tasks.

The participants will get paid after reporting the information. The payment is decided based on the reports across all T tasks.

Definition 2.1 (Multi-task peer prediction mechanism). *A multi-task peer prediction mechanism asks the participants to report their private information $\mathbf{r}_i^{(1:T)}$ for all T tasks. Then the payment to a participant i is decided based on all the reports, denoted by $p_i(\widetilde{\mathbf{r}}_1^{(1:T)}, \ldots, \widetilde{\mathbf{r}}_n^{(1:T)})$ when participant i's actual report is $\widetilde{\mathbf{r}}_i^{(1:T)}$.*

In general, the payment rule for multiple tasks $\mathbf{p}(\widetilde{\mathbf{r}}_1^{(1:T)}, \ldots, \widetilde{\mathbf{r}}_n^{(1:T)})$ can be very complicated. But in practice, we would prefer mechanisms that have succinct payment rules. In this work, we will consider a class of mechanisms that we call the *scoring mechanisms*.

Definition 2.2 (Scoring mechanisms). *A* scoring mechanism *assigns a payment to each of participant i's report $\widetilde{\mathbf{r}}_i^{(t)}$ by comparing it with other participants' report $\widetilde{\mathbf{r}}_{-i}$. Formally, a scoring mechanism uses a payment rule that can be represented as follows:*

$$p_i(\widetilde{\mathbf{r}}_i^{(1:T)}, \widetilde{\mathbf{r}}_{-i}^{(1:T)}) = \sum_{t=1}^{T} p_i^{(t)}(\widetilde{\mathbf{r}}_i^{(t)}, \widetilde{\mathbf{r}}_{-i}^{(1:T)}). \tag{1}$$

The key feature of scoring mechanisms is that the payment is decided separately for each of participant's reports for different tasks. To our knowledge, all the existing mechanisms that achieve strict truthfulness belong to scoring mechanisms.

The participants' goal is to maximize their own expected payoff. In multiple task peer prediction literature, it is always assumed that an agent's reporting strategy for task t only depends on his signal for that task $s_i^{(t)}$ but not the signals for other tasks $s_i^{(-t)}$. So we define a randomized strategy of agent i as follows.

Definition 2.3. *A strategy of agent i for a single task is a mapping $\sigma_i : \mathcal{S}_i \to \Delta(\mathcal{R}_i)$ that maps his observed signal s_i for that task into a distribution of reports, so that when agent i adopts strategy σ_i, he randomly draw a report $\widetilde{\mathbf{r}}_i$ according to $\sigma_i(s_i)$ when the observed signal is s_i.*

We denote agent i's strategy for task t by $\sigma_i^{(t)}$. Then we say a multiple-task mechanism is strictly truthful if truthfully reporting $\mathbf{r}_i^{(t)}$ for all tasks is a strict BNE.

Definition 2.4 (Strict truthfulness). *A payment rule $\mathbf{p}(\widetilde{\mathbf{r}}^{(1:T)})$ is strictly truthful for a distribution $\mu(\omega, s_1, \ldots, s_n)$ if, assuming that the participants know $\mu(\omega, s_1, \ldots, s_n)$, truthfully reporting $\mathbf{r}_i^{(1:T)}$ is a strict BNE, i.e., for any non-truthful strategy $\sigma_i^{(1:T)}$ with $\sigma_i^{(t)}(s_i^{(t)}) \neq f_{i,\mu}(s_i^{(t)})$ for some $s_i^{(t)}$,*

$$\mathbb{E}_\mu[p_i(\mathbf{R}_i^{(1:T)}, \mathbf{R}_{-i}^{(1:T)})] > \mathbb{E}_\mu[p_i(\sigma_i^{(1:T)}(\mathbf{S}_i^{(1:T)}), \mathbf{R}_{-i}^{(1:T)})].$$

Here we abuse the notation that $\sigma_i^{(1:T)}(\mathbf{S}_i^{(1:T)})$ represents the vector $(\sigma_i^{(1)}(S_i^{(1)}), \ldots, \sigma_i^{(T)}(S_i^{(T)}))$.

Another commonly used assumption is that the participants use consistent strategies across all tasks. This assumption is usually justified by that the designer can randomly shuffle the tasks so that the participants cannot distinguish the tasks (e.g. see [22]).

Assumption 2.5. *The participants use consistent strategies over all the tasks, i.e.,* $\sigma_i^{(1)} = \sigma_i^{(2)} = \cdots = \sigma_i^{(T)}$ *for any participant* i.

Then we define strict truthfulness under consistent strategies as follows.

Definition 2.6 (Strict truthfulness under consistent strategies). *A payment rule* $\mathbf{p}(\vec{\mathbf{r}}^{(1:T)})$ *is strictly truthful for a distribution* $\mu(\omega, s_1, \ldots, s_n)$ *if, assuming that the participants know the underlying distribution* $\mu(\omega, s_1, \ldots, s_n)$*, truthfully reporting* $\mathbf{r}_i^{(1:T)}$ *is a strict BNE, i.e., for all* i*, strategy* σ_i *with* $\sigma_i(s_i) \neq f_{i,\mu}(s_i)$ *for some* s_i*,*

$$\mathbb{E}_\mu[p_i(\mathbf{R}_i^{(1:T)}, \mathbf{R}_{-i}^{(1:T)})] > \mathbb{E}_\mu[p_i(\sigma_i(\mathbf{S}_i^{(1:T)}), \mathbf{R}_{-i}^{(1:T)})].$$

Here we abuse the notation that $\sigma_i(\mathbf{S}_i^{(1:T)})$ *represents the vector* $(\sigma_i(S_i^{(1)}), \ldots, \sigma_i(S_i^{(T)}))$.

The definition of truthfulness is with respect to the single true underlying distribution μ. For a designer who does not know μ but only knows that $\mu \in M$, we say that $\mathbf{r}$ is elicitable if there exists a mechanism that is strictly truthful for any $\mu \in M$.

Definition 2.7 (Elicitability). *A multi-task peer prediction problem* $\langle \mathbf{f}, M \rangle$ *with* $\mathbf{f} = \{f_{i,\mu}\}_{(i,\mu)}$ *is elicitable if there exists a payment rule* $\mathbf{p}(\vec{\mathbf{r}}^{(1:T)})$ *that is strictly truthful for any possible underlying distribution* $\mu \in M$.

If we assume that the participants use consistent strategies, the the definition of elicitability only requires strict truthfulness under consistent strategies.

Definition 2.8 (Elicitability under consistent strategies). *A multi-task peer prediction problem* $\langle \mathbf{f}, M \rangle$ *with* $\mathbf{f} = \{f_{i,\mu}\}_{(i,\mu)}$ *is elicitable under consistent strategies if there exists a payment rule* $\mathbf{p}(\vec{\mathbf{r}}^{(1:T)})$ *that is strictly truthful under consistent strategies for any possible underlying distribution* $\mu \in M$.

In this work, we mainly focus on multi-task peer prediction. Single-task peer prediction can be seen as a special case with $T = 1$.

3 PRELIMINARY

In this section, we review a few important results from previous works and add some minor findings. First, we give the definition of *power diagrams* and restate the characterization of elicitable problems in the single-task setting. Next, we discuss the correlated agreement mechanism from [22] and the mechanism that elicits participants' prediction about the state from [11].

3.1 Characterization for single-task elicitability

Frongillo and Witkowski [5] characterized the elicitability of the single-task problem using a geometric approach from the literature on property elicitation [13, 14]. The basic idea is that the agents' possible posterior beliefs need to fall into the correct regions. The regions are described by *power diagrams*.

Definition 3.1 ([5, 13]). *A* power diagram *in dimension* m *with* K *cells is a partitioning of* Δ_m *into* K *sets called* cells*, defined by a collection of* K *m-dimensional points* $\{\mathbf{v}^k \in \mathbb{R}^m : k \in [K]\}$ *called* sites *with associated* weights $\{w^k \in \mathbb{R} : k \in [K]\}$*, given by*

$$\mathit{cell}(\mathbf{v}^k) = \left\{ \mathbf{u} \in \mathbb{R}^m : \{k\} = \arg\min_{x \in [K]} \langle \mathbf{u}, \mathbf{v}^x \rangle - w^x \right\}.$$

Here $\langle \mathbf{u}, \mathbf{v}^x \rangle$ *represents the inner product of the two vectors. We call* $\langle \mathbf{u}, \mathbf{v}^x \rangle - w^x$ *the* power distance *from* $\mathbf{u}$ *to site* $\mathbf{v}^x$*; thus, for every point* $\mathbf{u}$ *in cell*$(\mathbf{v}^k)$*, it holds that* $\mathbf{v}^k$ *is closer to* $\mathbf{u}$ *in power distance than any other site* $\mathbf{v}^x$*.*

Here we use a definition of power diagrams that is slightly different from the original definition in [5], which used $\|\mathbf{u} - \mathbf{v}^x\|^2 - w^x$ as the power distance. The two definitions are equivalent and can be transformed into one another by changing the value of $\mathbf{w}^x$. For our multi-task problem, we find it more convenient to use $\langle \mathbf{u}, \mathbf{v}^x \rangle - w^x$ as the power distance.

To give the characterization in Frongillo and Witkowski [5], we define the following. Let $Q_i(\mathbf{r}_i)$ be the set of agent i's possible posterior belief of $\mathbf{r}_{-i}$ when his truthful report is $\mathbf{r}_i$.

$$Q_i(\mathbf{r}_i) = \{\mu(\mathbf{r}_{-i}|s_i) : \mu \in M,\ s_i \text{ satisfies } f_{i,\mu}(s_i) = \mathbf{r}_i\}. \tag{2}$$

Theorem 3.2. *A single-task elicitability problem* $\langle \mathbf{f}, M \rangle$ *is elicitable if and only if for each agent* i*, there exists a power diagram in dimension* $|\mathcal{R}_{-i}|$ *with* $|\mathcal{R}_i|$ *cells defined by* $\{\mathbf{v}^{\mathbf{r}_i} \in \mathbb{R}^{|\mathcal{R}_{-i}|} : \mathbf{r}_i \in \mathcal{R}_i\}$ *and* $\{w^{\mathbf{r}_i} \in \mathbb{R} : \mathbf{r}_i \in \mathcal{R}_i\}$*, such that each* $Q_i(\mathbf{r}_i)$ *falls into a distinct cell,*

$$Q_i(\mathbf{r}_i) \subseteq \textit{cell}(\mathbf{v}^{\mathbf{r}_i}),\ \textit{for all } \mathbf{r}_i \in \mathcal{R}_i.$$

The theorem gives the necessary and sufficient condition for a single-task problem to be elicitable. This result is given by Frongillo and Witkowski [5] in the setting where the mechanism is minimal, i.e., the agents are asked to directly reveal their signals $\mathbf{r}_i = s_i$ and the support of the signals is finite. When the reports are arbitrary functions, it could be difficult to analyze the space $\Delta_{|\mathcal{R}_{-i}|}$, e.g. when the support of the reports is continuous.

We find the following proposition to be useful in our exposition, and include a proof in the appendix. Note that the result can be implied by Theorem 3.2 when report space is discrete, and we show it still holds when the support of $\mathbf{r}_{-i}$ is continuous.

Proposition 3.3 (Robust stochastic relevancy). *A single-task problem* $\langle \mathbf{f}, M \rangle$ *is elicitable only if for any* i *and* $\mathbf{r}_i \neq \mathbf{r}_i' \in \mathcal{R}_i$*,* $Q_i(\mathbf{r}_i) \cap Q_i(\mathbf{r}_i') = \emptyset$*.*

If f is identity, i.e. $f(s_i) = s_i$, and $M = \{\mu\}$, the above condition implies stochastic relevancy. Thus, we call the above characterization *robust stochastic relevancy*.

Finally, if the designer's knowledge about the information structure is accurate enough, it is possible to design a strictly truthful mechanism. We defer the details to Appendix B.

3.2 Correlated agreement mechanism

Shnayder et al. [22] proposed the *correlated agreement mechanism* for multi-task peer prediction. They considered the design of *minimal mechanisms*, that is, mechanisms that ask the agents to directly report their signals, so we have

$$\mathbf{r}_i = s_i,\ \forall i.$$

The correlated agreement mechanism only needs two participants $n = 2$ and two tasks $T = 2$. The mechanism requires the designer to know the correlation structure of signals, but not the full signal distribution. More specifically, define the *Delta matrix* Δ to be a $|\mathcal{S}_1| \times |\mathcal{S}_2|$ matrix with entry in row s_1 and column s_2 equal to

$$\Delta[s_1, s_2] = \mu(s_1, s_2) - \mu(s_1)\mu(s_2),$$

where $\mu(s_1, s_2)$ is the joint distribution of the two participants' signals, and $\mu(s_1), \mu(s_2)$ are the marginal distributions of s_1 and s_2 respectively. The Delta matrix describes the correlation between different realized signal values. If an entry $\Delta[s_1, s_2] > 0$, then we have $\mu(s_1|s_2) > \mu(s_1)$ and $\mu(s_2|s_1) > \mu(s_2)$, which means that seeing s_2 will increase participant 2's belief about seeing s_1, and seeing s_1

will increase participant 1's belief about seeing s_2, so the signal realizations s_1 and s_2 are positively correlated. To ensure strict truthfulness, the CA mechanism requires the designer to know the sign of each entry of the Delta matrix, denoted by Sign(Δ), which means that the designer needs to know for each pair of signal realizations whether they are positively correlated or negatively correlated. The payment of the CA mechanism is designed as follows.

Definition 3.4 (Correlated agreement mechanism [22]). *The correlated agreement mechanism asks the two participants to report their signals for two tasks. The payment to participant $i \in \{1,2\}$ for task $t \in \{1,2\}$ is*

$$p_i^{(t)}(\mathbf{s}^{(1:T)}) = Sign(\Delta[s_i^{(t)}, s_{-i}^{(t)}]) - Sign(\Delta[s_i^{(t)}, s_{-i}^{(-t)}]),$$

where $Sign(\Delta[x,y])$ represents the sign of the entry in row x and column y of matrix Δ.

The above definition is slightly different from the original definition in [22], but they are equivalent in the sense of elicitability. The CA mechanism will be strictly truthful if the following condition holds.

Theorem 3.5 (Shnayder et al. [22]). *If the matrix $Sign(\Delta)$ does not have two identical rows and it does not have two identical columns, the correlated agreement mechanism is strictly truthful.*

Shnayder et al. [22] also discussed other incentive properties. In this work, we only consider the strict truthfulness defined in Definition 2.6, which is equivalent to the definition of strict properness (Definition 2.6) in their work.

3.3 Elicit predictions

Kong and Schoenebeck [11] proposed a mechanism that elicits the participants' posterior about the state $\mu(\omega|s_i)$ in both single-task and multiple-task settings, when the prior $\mu(\omega)$ is known to the designer and the participants' signals are independent conditioning on the state ω, that is,

$$\mu(\omega, s_1, \dots, s_n) = \mu(\omega)\mu(s_1|\omega)\cdots\mu(s_n|\omega),\ \forall \omega, s_1, \dots, s_n.$$

Chen et al. [2] further give a sufficient condition for the mechanisms to be strictly truthful. For a distribution μ, define P_i^μ to be a $|\mathcal{S}_{-i}| \times |\Omega|$ matrix with entry in row $\mathbf{s}_{-i}$ and column ω equal to $\mu(\mathbf{s}_{-i}|\omega)$. Then we have the follows.

Lemma 3.6 ([2, 11]). *Consider a single-task/multi-task problem $\langle \mathbf{f}, M\rangle$ with $\mathbf{r}_i = \mu(\omega|s_i)$. Suppose the prior $\mu(\omega)$ is known to the designer and the participants' signals are independent conditioning on the state ω. If we further have $rank(P_i^\mu) = |\Omega|$ for all $\mu \in M$ and $i \in [n]$, then $\mathbf{r}_i = \mu(\omega|s_i)$ is elicitable by the following payment rule for every single task*

$$p_i(\mathbf{r}_i, \mathbf{r}_{-i}) = \log\left(\sum_{\omega\in\Omega} \frac{\mathbf{r}_i(\omega) g(\mathbf{r}_{-i}, \omega)}{\mu(\omega)}\right),$$

with

$$g(\mathbf{r}_{-i}, \omega) = \frac{1}{A(\mathbf{r}_{-i})} \cdot \frac{\prod_{j\neq i} \mathbf{r}_j(\omega)}{\mu(\omega)^{n-1}},$$

where $A(\mathbf{r}_{-i})$ is a normalization term so that $\sum_\omega g(\mathbf{r}_{-i}, \omega) = 1$.

4 MULTIPLE-TASK ELICITABILITY

As introduced in Section 3.1, if the designer only collects reports for one task, Frongillo and Witkowski [5] showed that a single-task peer prediction problem is elicitable if and only if each agent's posteriors can be fitted into a power diagram. But when the designer has multiple i.i.d.

tasks, it is possible for the designer to exploit the similarity between the tasks and elicit information that is not elicitable in the single-task framework. For example, Shnayder et al. [22] showed that the Dasgupta-Ghosh mechanism [3] can elicit $r_i = s_i$ when $n = 2$ and the signals are categorical, which means that when an agent sees a signal, all other signals become less likely than their prior probability, i.e.,

$$\mu(s_2 = y|s_1 = x) < \mu(s_2 = y), \forall x \neq y.$$

But the categorical condition clearly does not guarantee the *robust stochastic relevance* (Corollary 3.3) without the knowledge about the marginal distribution, which is the necessary condition for a problem to be elicitable in the single-task framework. Therefore we need stronger conditions for the elicitability of multi-task problems.

In this section, we first give a necessary and sufficient condition for a multi-task problem to be elicitable if the designer only uses *scoring mechanisms*. This characterization holds when $\mathcal{R}_i$'s are finite sets, i.e., there are finitely many possible values of a participant's report. We show how to use our characterization in the setting of the CA mechanisms (Section 3.2). For the general case when $\mathcal{R}_i$'s can be infinitely large, we provide necessary conditions for a multi-task problem to be elicitable. One of the necessary conditions will be the key tool that we use to obtain the results in Section 5. Second, for general mechanisms, we give a necessary condition for a multi-task problem to be elicitable, assuming that the participants use consistent strategies.

4.1 Scoring mechanisms

The challenge of studying the elicitability of multi-task problems is largely due to the complexity of the payment rule $\mathbf{p}(\mathbf{r}^{(1:T)})$. The payment rule $\mathbf{p}(\mathbf{r}^{(1:T)})$ can potentially be an extremely complicated function, which may not even be efficiently computable as $\mathbf{r}^{(1:T)}$ has exponentially many possible values. But in practice, such payments are unlikely to be appealing because of the implementation difficulty as well as the lack of transparency. So in this section, we restrict our attention to a smaller class of mechanisms: the scoring mechanisms (Definition 2.2). The scoring mechanisms pay each of a participant's reports separately. The payment to a report is decided by comparing it with other participants' reports across all the tasks. To our knowledge, all of the existing strictly truthful mechanisms belong to scoring mechanisms.

Before giving our main results, we first show that when we focus on the elicitability by scoring mechanisms, it does not matter whether the participants' strategies will be consistent or not.

Theorem 4.1. *When the designer only uses scoring mechanisms, a multi-task problem $\langle \mathbf{f}, M \rangle$ is elicitable if and only if it is elicitable under consistent strategies.*

The proof of Theorem 4.1 can be found in Appendix C. In the rest of this section, we just assume that the participants can use non-consistent strategies. The characterization of elicitability under consistent strategies is just the same.

4.1.1 Characterization of elicitable multi-task problems. We now give the characterization of elicitable multi-task problems when all $\mathcal{R}_i$'s are finite. Our first observation is that if a multi-task problem $\langle \mathbf{f}, M \rangle$ is elicitable by a scoring mechanism, then for each participant i, his posteriors should fall into correct cells of a power diagram *for any given marginal distribution of other participants' truthful reports*. To show this formally, we introduce some notations. Consider an agent i and a given marginal distribution of other agents' truthful reports for a single task $\mu(\mathbf{r}_{-i})$. Define $M_{\mu(\mathbf{r}_{-i})}$ to be the set of all distributions that has marginal distribution of $\mathbf{r}_{-i}$ equal to $\mu(\mathbf{r}_{-i})$,

$$M_{\mu(\mathbf{r}_{-i})} = \{\mu' \in M : \mu'(\mathbf{r}_{-i}) = \mu(\mathbf{r}_{-i})\}$$

Also define $Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))$ to be the set of participant i's possible posteriors about $\mathbf{r}_{-i}$ when participant i's truthful report is $\mathbf{r}_i$ and the marginal distribution of $\mathbf{r}_{-i}$ is $\mu(\mathbf{r}_{-i})$,

$$Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i})) = \{\mu'(\mathbf{r}_{-i}|s_i) : \mu' \in M_{\mu(\mathbf{r}_{-i})},\ s_i \text{ satisfies } f_{i,\mu}(s_i) = \mathbf{r}_i\}.$$

Then a necessary condition for a multi-task problem $\langle \mathbf{f}, M\rangle$ to be elicitable is that for any given $\mu(\mathbf{r}_{-i})$, there exists a power diagram that divides $Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))$ for different $\mathbf{r}_i$ into different cells. We call this the power diagram constraint for given marginal distributions.

Definition 4.2. *A multi-task problem $\langle \mathbf{f}, M\rangle$ with finite-size $\mathcal{R}_i$'s satisfies the* power diagram constraint for given marginal distributions *if for all i and $\mu(\mathbf{r}_{-i})$, $\{Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))\}_{\mathbf{r}_i \in \mathcal{R}_i}$ can be fitted into a power diagram, which means that there exists a power diagram in dimension $|\mathcal{R}_{-i}|$ with $|\mathcal{R}_i|$ cells defined by $\{\mathbf{v}^{\mathbf{r}_i} \in \mathbb{R}^{|\mathcal{R}_{-i}|} : \mathbf{r}_i \in \mathcal{R}_i\}$ with associated weights $\{w^{\mathbf{r}_i} \in \mathbb{R} : \mathbf{r}_i \in \mathcal{R}_i\}$, such that each $Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))$ falls into a distinct cell,*

$$Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i})) \subseteq \mathit{cell}(\mathbf{v}^{\mathbf{r}_i}),\ \text{for all } \mathbf{r}_i \in \mathcal{R}_i.$$

Note that checking this condition does not require the designer to know the actual marginal distribution $\mu(\mathbf{r}_{-i})$. The condition means that for any given marginal distribution $\mu(\mathbf{r}_{-i})$, if the designer restrict the possible underlying distributions to the ones that has marginal distribution of $\mathbf{r}_{-i}$ equal to the given $\mu(\mathbf{r}_{-i})$, it should be possible to fit $\{Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))\}_{\mathbf{r}_i \in \mathcal{R}_i}$ into a power diagram.

Below we provide an example to illustrate the sets $Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))$ and how they can be fitted into power diagrams.

Example 4.3. *Consider the following problem instance $\langle \mathbf{f}, M\rangle$. Suppose there are two agents $n = 2$, two tasks $T = 2$ and the report and signal spaces are $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{R}_1 = \mathcal{R}_2 = \{0, 1, 2\}$. The designer asks the participants to directly report their signals $r_i = f_{i,\mu}(s_i) = s_i$. The set of possible distributions M is the set of all distributions that have the sign of the Delta matrix Sign(Δ) (defined in Section 3.2) equal to*

$$\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

This means that for a participant i, seeing a signal realization $s_i^* \in \{0, 1, 2\}$ will increase the probability that the other participant also observes the same signal realization s_i^* but decrease the probability that the other participant observes a different signal realization. Given such M and $\mathcal{R}_i = \{0, 1, 2\}$, we use simplex plot on $\Delta(\{0, 1, 2\})$ to illustrate the sets $Q_i(r_i, \mu(r_{-i}))$. In Figure 1, each point on the simplex plot represents a distribution on $\{0, 1, 2\}$. We choose two marginal distributions, $\mu(r_{-i}) = (1/3, 1/3, 1/3)$ on the left and $\mu(r_{-i}) = (0.2, 0.3, 0.5)$ on the right. The colored areas are the set $\{Q_i(r_i, \mu(r_{-i}))\}_{r_i \in \mathcal{R}_i}$, and the dashed lines are the boundary of the cells of a power diagram with certain sites and weights. The exact definition of the sites and the weights will be given in (9) *later in Section 4.1.2 when we discuss the application of our results.*

Now assume that $\langle \mathbf{f}, M\rangle$ satisfies the power diagram constraint for given marginal distributions. For any report $\mathbf{r}_i \in \mathcal{R}_i$, denote the site $\mathbf{v}^{\mathbf{r}_i}$ for $\mu(\mathbf{r}_{-i})$ by $\mathbf{v}^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i}))$ and denote the associated weight by $w^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i}))$. Here we abuse the notation and consider $w^{r_i} : \Delta(\mathcal{R}_{-i}) \to \mathbb{R}$ and $\mathbf{v}^{\mathbf{r}_i} : \Delta(\mathcal{R}_{-i}) \to \mathbb{R}^{|\mathcal{R}_{-i}|}$ as functions of the marginal distribution $\mu(\mathbf{r}_{-i})$. Our second observation is that the sites of the power diagrams $\mathbf{v}^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i}))$ and the associated weights $w^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i}))$ need to be affine functions of the marginal distribution of other participants' reports for $T-1$ tasks $\mu(\mathbf{r}_{-i}^{(1:T-1)})$.[2] To be more specific, let $\mathbf{u}$ be the length-$|\mathcal{R}_{-i}|$ vector that represents the distribution $\mu(\mathbf{r}_{-i})$. Then $\mu(\mathbf{r}_{-i}^{(1:T-1)})$ can

[2]Because the tasks are i.i.d., the distribution for $T-1$ tasks $\mu(\mathbf{r}_{-i}^{(1:T-1)})$ can be generated by the distribution for a single task $\mu(\mathbf{r}_{-i})$.

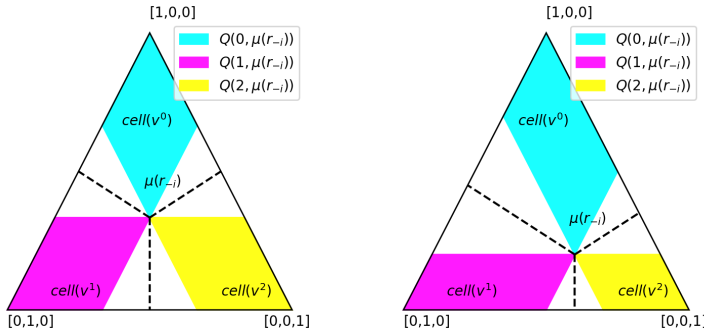

Fig. 1. With various marginal distributions $\mu(r_{-i})$ and $r_i = 0, 1, 2$, the simplex plots illustrate the sets $\{Q_i(r_i, \mu(r_{-i}))\}_{r_i \in \mathcal{R}_i}$ (colored areas) defined in Example 4.3 and the cells of power diagrams cell($\mathbf{v}^{r_i}$) (with dashed boundary) that separate them. In our example, the point at the intersection of the dashed lines is the marginal distributions $\mu(r_{-i})$.

be represented by the $(T-1)$-th tensor power of the vector, $\mathbf{u}^{\otimes(T-1)}$. Our observation is that both $\mathbf{v}^{\mathbf{r}_i}(\mathbf{u})$ and $w^{\mathbf{r}_i}(\mathbf{u})$ need to be affine functions of $\mathbf{u}^{\otimes(T-1)}$. Moreover, when such power diagrams and affine functions exist, we can find a mechanism that elicits $\mathbf{r}$.

Theorem 4.4. *A multi-task problem $\langle \mathbf{f}, M \rangle$ with finite-size $\mathcal{R}_i$'s is elicitable by scoring mechanisms if and only if*

(1) It satisfies the power diagram constraint for given marginal distributions. Let $\mathbf{u}$ be a vector that represents a marginal distribution $\mu(\mathbf{r}_{-i})$ and denote by $\mathbf{v}^{\mathbf{r}_i}(\mathbf{u})$ and $w^{\mathbf{r}_i}(\mathbf{u})$ the sites and the weights of the power diagram for the marginal distribution $\mathbf{u}$.

(2) Furthermore, for every $i \in [n]$ and every $\mathbf{r}_i \in \mathcal{R}_i$, there exist a matrix $\mathbf{D}_{\mathbf{r}_i}$ with $|\mathcal{R}_{-i}|$ rows and $|\mathcal{R}_{-i}|^{T-1}$ columns and a vector $\mathbf{e}_{\mathbf{r}_i} \in \mathbb{R}^{|\mathcal{R}_{-i}|}$ with

$$\mathbf{v}^{\mathbf{r}_i}(\mathbf{u}) = \mathbf{D}_{\mathbf{r}_i} \cdot \mathbf{u}^{\otimes(T-1)} + \mathbf{e}_{\mathbf{r}_i}, \quad \text{for all possible } \mathbf{u},$$

and there exists a vector $\mathbf{h}_{\mathbf{r}_i}$ with length $|\mathcal{R}_{-i}|^{T-1}$ such that

$$w^{\mathbf{r}_i}(\mathbf{u}) = \mathbf{h}_{\mathbf{r}_i}^\top \cdot \mathbf{u}^{\otimes(T-1)}, \quad \text{for all possible } \mathbf{u}.$$

Moreover, if such power diagrams and affine functions exist, we can find a mechanism that is strictly truthful for any $\mu \in M$ with payments defined by the entries of $\mathbf{D}, \mathbf{e}$ and $\mathbf{h}$ as

$$p_i^{(t)}(\mathbf{r}_i, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}) = -\mathbf{D}_{\mathbf{r}_i}[\mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}] - \mathbf{e}_{\mathbf{r}_i}[\mathbf{r}_{-i}^{(t)}] + \mathbf{h}_{\mathbf{r}_i}[\mathbf{r}_{-i}^{(-t)}], \quad \forall \mathbf{r}_i \in \mathcal{R}_i,\ \mathbf{r}_{-i}^{(t)},\ \mathbf{r}_{-i}^{(-t)}. \tag{3}$$

We want to point out that the parameters $\mathbf{e}, \mathbf{h}$ in the theorem can actually be merged into matrix $\mathbf{D}$ so that $\mathbf{D}_{\mathbf{r}_i} = \mathbf{e}_{\mathbf{r}_i} \cdot \mathbf{1}^\top + \mathbf{1} \cdot \mathbf{h}_{\mathbf{r}_i}^\top$. We use this form because in some applications, it is more convenient to separate $\mathbf{e}$ and $\mathbf{h}$ and have a payment in form (3), for example when we apply the theorem to the setting of the correlated agreement mechanism [22].

Proof of Theorem 4.4. We consider the general setting when the agents can use non-consistent strategy over different tasks.

We first prove the necessity of the condition. Suppose that the designer has a strictly truthful scoring mechanisms for a multi-task problem $\langle \mathbf{f}, M \rangle$, so that the payment has the form

$$p_i(\widetilde{\mathbf{r}}_i^{(1:T)}, \widetilde{\mathbf{r}}_{-i}^{(1:T)}) = \sum_{t=1}^{T} p_i^{(t)}(\widetilde{\mathbf{r}}_i^{(t)}, \widetilde{\mathbf{r}}_{-i}). \tag{4}$$

When participant i decides the strategy for task t, his report for this task $\widetilde{\mathbf{r}}_i^{(t)}$ only affects the payment for this task $p_i^{(t)}(\widetilde{\mathbf{r}}_i^{(t)}, \widetilde{\mathbf{r}}_{-i})$. Therefore his best strategy for task t is the one that maximizes the expected payment for task t assuming that the other agents truthfully report, that is, $\mathbb{E}_\mu[p_i^{(t)}(\widetilde{\mathbf{R}}_i^{(t)}, \mathbf{R}_{-i})]$. This means that for each signal realization $s_i^{(t)}$, participant i should choose the report $\widetilde{\mathbf{r}}_i^{(t)}$ that will maximize the conditional expectation $\mathbb{E}_\mu[p_i^{(t)}(\widetilde{\mathbf{R}}_i^{(t)}, \mathbf{R}_{-i})|s_i^{(t)}]$. Since the tasks are independent, observing $s_i^{(t)}$ will only change agent i's belief about other people's reports for this task $\mathbf{r}_{-i}^{(t)}$, but not the reports for other tasks $\mathbf{r}_{-i}^{(-t)}$. So we can factor out $\mathbf{r}_{-i}^{(t)}$ in the conditional expectation as follows

$$\begin{aligned}
\mathbb{E}_\mu\left[p_i^{(t)}\left(\widetilde{\mathbf{R}}_i^{(t)}, \mathbf{R}_{-i}^{(t)}, \mathbf{R}_{-i}^{(-t)}\right)|s_i^{(t)}\right] &= \sum_{\mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}} \mu(\mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}|s_i^{(t)}) \cdot p_i^{(t)}\left(\widetilde{\mathbf{r}}_i^{(t)}, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}\right) \\
&= \sum_{\mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}} \mu(\mathbf{r}_{-i}^{(t)}|s_i^{(t)}) \cdot \mu(\mathbf{r}_{-i}^{(-t)}) \cdot p_i^{(t)}\left(\widetilde{\mathbf{r}}_i^{(t)}, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}\right) \\
&= \sum_{\mathbf{r}_{-i}^{(t)}} \mu(\mathbf{r}_{-i}^{(t)}|s_i^{(t)}) \cdot \sum_{\mathbf{r}_{-i}^{(-t)}} \mu(\mathbf{r}_{-i}^{(-t)}) \cdot p_i^{(t)}\left(\widetilde{\mathbf{r}}_i^{(t)}, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}\right). \qquad (5)
\end{aligned}$$

Now consider a given marginal distribution of other participants' report for one task, $\mu(\mathbf{r}_{-i})$. Define $\alpha_y(\mathbf{r}_{-i}^{(t)})$ to be the value of the second sum in (5) when participant i reports y, that is, $\widetilde{\mathbf{r}}_i^{(t)} = y$, and other participants' truthful report for task t is $\mathbf{r}_{-i}^{(t)}$,

$$\alpha_y(\mathbf{r}_{-i}^{(t)}) = \sum_{\mathbf{r}_{-i}^{(-t)}} \mu(\mathbf{r}_{-i}^{(-t)}) \cdot p_i^{(t)}\left(y, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}\right),$$

where the distribution $\mu(\mathbf{r}_{-i}^{(-t)})$ is generated according to the given marginal distribution $\mu(\mathbf{r}_{-i})$. Then participant i's expected payoff (for task t) conditioning on $s_i^{(t)}$ when reporting y is the inner product of the vectors that represent posterior belief $\mu(\mathbf{r}_{-i}^{(t)}|s_i^{(t)})$ and $\alpha_y(\mathbf{r}_{-i}^{(t)})$,

$$\mathbb{E}_\mu\left[p_i^{(t)}\left(y, \mathbf{R}_{-i}^{(t)}, \mathbf{R}_{-i}^{(-t)}\right)|s_i^{(t)}\right] = \sum_{\mathbf{r}_{-i}^{(t)}} \mu(\mathbf{r}_{-i}^{(t)}|s_i^{(t)}) \cdot \alpha_y(\mathbf{r}_{-i}^{(t)}) = \langle \boldsymbol{\mu}(\mathbf{r}_{-i}^{(t)}|s_i^{(t)}), \boldsymbol{\alpha}_y \rangle.$$

Here we abuse the notation to use $\boldsymbol{\mu}(\mathbf{r}_{-i}^{(t)}|s_i^{(t)})$ to represent a vector. As a result, by the definition of strict truthfulness, when agent i's truthful report is $\mathbf{r}_i$, his expected payoff (for task t) should be uniquely maximized when reporting $\mathbf{r}_i$, which means

$$\{\mathbf{r}_i\} = \arg\max_{y \in \mathcal{R}_i} \langle \boldsymbol{\mu}(\mathbf{r}_{-i}|s_i), \boldsymbol{\alpha}_y \rangle. \qquad (6)$$

We construct a power diagram by setting $\mathbf{v}^y = -\boldsymbol{\alpha}_y$ and $w^y = 0$. Then (6) is equivalent to

$$\mu(\mathbf{r}_{-i}|s_i) \in \text{cell}(\mathbf{v}^{\mathbf{r}_i}).$$

Therefore if $\langle \mathbf{f}, M \rangle$ is elicitable by scoring mechanisms, $\{Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))\}_{r_i \in \mathcal{R}_i}$ can be fitted into a power diagram. Furthermore, let matrix $\mathbf{D}_y$ be a $|\mathcal{R}_{-i}| \times |\mathcal{R}_{-i}|^{T-1}$ matrix with $p_i^{(t)}(y, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)})$ in row $\mathbf{r}_{-i}^{(t)}$ and column $\mathbf{r}_{-i}^{(-t)}$. Then by our construction of $\mathbf{v}^y$, we have

$$\mathbf{v}^y(\mu(\mathbf{r}_{-i})) = \alpha_y = \mathbf{D}_y \cdot \mu(\mathbf{r}_{-i}^{(-t)}).$$

Therefore it is necessary that there exists $\mathbf{v}^y(\mu(\mathbf{r}_{-i}))$ which is an affine function of $\mu(\mathbf{r}_{-i}^{(-t)})$.

Next, we prove the sufficiency of the condition. If we have power diagrams defined by $\{\mathbf{v}^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i})) \in \mathbb{R}^{|\mathcal{R}_{-i}|} : \mathbf{r}_i \in \mathcal{R}_i\}$ and $w(\mathbf{r}_i)$ such that for any given $\mu(\mathbf{r}_{-i})$, $Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))$ falls into a distinct cell

$$Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i})) \subseteq \text{cell}(\mathbf{v}^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i}))), \text{ for all } \mathbf{r}_i \in \mathcal{R}_i. \tag{7}$$

and for each $\mathbf{r}_i$, there exist a matrix $\mathbf{D}_{\mathbf{r}_i}$ with $|\mathcal{R}_{-i}|$ rows and $|\mathcal{R}_{-i}|^{T-1}$ columns and a vector $\mathbf{e}_{\mathbf{r}_i} \in \mathbb{R}^{|\mathcal{R}_{-i}|}$ such that

$$\mathbf{v}^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i})) = \mathbf{D}_{\mathbf{r}_i} \cdot \mu(\mathbf{r}_{-i}^{(-t)}) + \mathbf{e}_{\mathbf{r}_i},$$

and there exists a vector $\mathbf{h}$ with length $|\mathcal{R}_{-i}|^{T-1}$ such that

$$w^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i})) = \mathbf{h}_{\mathbf{r}_i}^\top \cdot \mu(\mathbf{r}_{-i}^{(-t)}).$$

Consider the payment rule (3) given by the theorem

$$p_i^{(t)}(\mathbf{r}_i, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}) = -\mathbf{D}_{\mathbf{r}_i}[\mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}] \ - \mathbf{e}_{\mathbf{r}_i}[\mathbf{r}_{-i}^{(t)}] + \mathbf{h}_{\mathbf{r}_i}[\mathbf{r}_{-i}^{(-t)}].$$

Then for any underlying distribution μ, suppose participant i observes $s_i^{(t)}$ for task t, if participant i reports $y \in \mathcal{R}_i$, as shown in (5), his expected payment for task t is

$$\begin{aligned}
\mathbb{E}_\mu\left[p_i^{(t)}(y, \mathbf{R}_{-i}^{(t)}, \mathbf{R}_{-i}^{(-t)})|s_i^{(t)}\right] &= \sum_{\mathbf{r}_{-i}^{(t)}} \mu(\mathbf{r}_{-i}^{(t)}|s_i^{(t)}) \cdot \sum_{\mathbf{r}_{-i}^{(-t)}} \mu(\mathbf{r}_{-i}^{(-t)}) \cdot p_i^{(t)}\left(y, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}\right) \\
&= -\langle \boldsymbol{\mu}(\mathbf{r}_{-i}^{(t)}|s_i^{(t)}), \mathbf{D}_y \cdot \boldsymbol{\mu}(\mathbf{r}_{-i}^{(-t)}) + \mathbf{e}_y \rangle + w^y(\mu(\mathbf{r}_{-i})) \\
&= -\langle \boldsymbol{\mu}(\mathbf{r}_{-i}^{(t)}|s_i^{(t)}), \mathbf{v}^y(\mu(\mathbf{r}_{-i})) \rangle + w^y(\mu(\mathbf{r}_{-i})).
\end{aligned}$$

Consequently, the payment rule will be strictly truthful because by the condition (7) we have

$$\mu(\mathbf{r}_{-i}|s_i) \in Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i})) \subseteq \text{cell}(\mathbf{v}^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i}))),$$

which by the definition of $\text{cell}(\mathbf{v}^{\mathbf{r}_i}(\mu(\mathbf{r}_{-i})))$ (Definition 3.1) means that

$$\begin{aligned}
\{\mathbf{r}_i\} &= \arg\min_{y \in \mathcal{R}_i} \langle \mathbf{v}^y(\mu(\mathbf{r}_{-i})), \boldsymbol{\mu}(\mathbf{r}_{-i}|s_i) \rangle - w^y(\mu(\mathbf{r}_{-i})) \\
&= \arg\min_{y \in \mathcal{R}_i} -E_\mu\left[p_i^{(t)}(y, \mathbf{R}_{-i}^{(t)}, \mathbf{R}_{-i}^{(-t)})|s_i^{(t)}\right] \\
&= \arg\max_{y \in \mathcal{R}_i} \mathbb{E}_\mu\left[p_i^{(t)}(y, \mathbf{R}_{-i}^{(t)}, \mathbf{R}_{-i}^{(-t)})|s_i^{(t)}\right].
\end{aligned}$$

□

Theorem 4.4 applies to the general scoring mechanisms whose $p_i^{(t)}(\widetilde{\mathbf{r}}_i^{(t)}, \widetilde{\mathbf{r}}_{-i}^{(1:T)})$ can be an arbitrary function of $\widetilde{\mathbf{r}}_i^{(t)}$ and $\widetilde{\mathbf{r}}_{-i}^{(1:T)}$. Since $\widetilde{\mathbf{r}}_{-i}^{(1:T)}$ has exponentially many possible values, the payment $p_i^{(t)}(\widetilde{\mathbf{r}}_i^{(t)}, \widetilde{\mathbf{r}}_{-i}^{(1:T)})$ of an arbitrary scoring mechanism can possibly be hard to compute. In practice, for computational reasons, the designer may only want to decide payments based on a function of $\widetilde{\mathbf{r}}_{-i}^{(1:T)}$, for example, a sufficient statistic of $\widetilde{\mathbf{r}}_{-i}^{(1:T)}$. In this case, we can have a simplified version of Theorem 4.4 with smaller $\mathbf{D}_{\mathbf{r}_i}$ and $\mathbf{h}_{\mathbf{r}_i}$, which will have polynomial sizes if the function of $\widetilde{\mathbf{r}}_{-i}^{(1:T)}$ has polynomially many possible values. We defer the details to Appendix C.2.

Our characterization can be used in two ways.

4.1.2 The sufficiency of the condition. Given a problem instance $\langle \mathbf{f}, M \rangle$, if we are able to find power diagrams as described in Theorem 4.4, then the payment rule (3) will guarantee strict truthfulness. To give an example, we apply Theorem 4.4 to the setting of the correlated agreement mechanism (Theorem 3.5). We show how to find power diagrams as described in Theorem 4.4 when the designer knows the sign of the Delta matrix Sign(Δ) and the matrix Sign(Δ) does not have two identical rows or two identical columns. We will find that the payment rule (3) induced by the power diagrams is just the payment rule of the CA mechanism.

Consider the setting for the CA mechanism in which $n = 2, T = 2$ and $r_i = s_i$ for all i. If the designer knows Sign(Δ) and Sign(Δ) does not contain two identical rows or columns, define $\text{Sign}(\Delta[s_1, :]) \in \{-1, +1\}^{|\mathcal{S}_2|}$ to be the sign of the s_1-th row of Δ. Our goal is to find power diagrams that will separate a participant i's possible posteriors of the other participant's report $\mu(s_2|s_1)$ for different signal realizations s_1, for any given marginal distribution $\mu(s_2)$. It may not be straightforward if we directly try to separate possible $\mu(s_2|s_1)$ for different s_1, but we can easily find power diagrams that separate $\mu(s_2|s_1) - \mu(s_2)$ for different s_1. By the definition of the Delta matrix Δ, $\mu(s_2|s_1) - \mu(s_2)$ is just the s_1-th row of Δ. If the designer knows the sign of Δ, we can just define the sites of a power diagram to be the rows of −Sign(Δ), that is, define

$$\widetilde{\mathbf{v}}^{s_1} = -\text{Sign}(\Delta[s_1, :]), \quad \widetilde{w}^{s_1} = 0, \quad \forall s_1 \in \mathcal{S}_1.$$

Then for a signal realization $s_1 \in \mathcal{S}_1$ and for any possible $\mu(s_2|s_1) - \mu(s_2)$, the site $\widetilde{\mathbf{v}}^{s_1}$ will be the closest to $\mu(s_2|s_1) - \mu(s_2)$ among all the sites in terms of power distance, that is,

$$\begin{aligned} \{s_1\} &= \arg\min_{y \in \mathcal{S}_1} \langle -\text{Sign}(\Delta[y, :]), \boldsymbol{\mu}(s_2|s_1) - \boldsymbol{\mu}(s_2) \rangle \\ &= \arg\min_{y \in \mathcal{S}_1} \langle -\text{Sign}(\Delta[y, :]), \boldsymbol{\mu}(s_2|s_1) \rangle + \langle \text{Sign}(\Delta[y, :]), \boldsymbol{\mu}(s_2) \rangle. \end{aligned} \tag{8}$$

Here we abuse the notation of $\boldsymbol{\mu}(s_2|s_1)$ and $\boldsymbol{\mu}(s_2)$ to denote the vectors that represent the distribution. This can be converted into power diagrams that will separate possible $\mu(s_2|s_1)$ for different s_1. From (8), we can define

$$\mathbf{v}^{s_1}(\mu(\mathbf{s}_2)) = -\text{Sign}(\Delta[s_1, :]), \text{ and } w^{s_1}(\mu(\mathbf{s}_2)) = -\langle \text{Sign}(\Delta[s_1, :]), \boldsymbol{\mu}(s_2) \rangle. \tag{9}$$

Fortunately, both $\mathbf{v}^{s_1}(\mu(\mathbf{s}_2))$ and $w^{s_1}(\mu(\mathbf{s}_2))$ are affine in $\mu(s_2)$, with parameters

$$\mathbf{D}_{s_1} = \mathbf{0}, \quad \mathbf{e}_{s_1} = -\text{Sign}(\Delta[s_1, :]), \quad \mathbf{h}_{s_1} = -\text{Sign}(\Delta[s_1, :]).$$

Then the payment defined by Theorem 4.4

$$p_i^{(t)}(r_i, \mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}) = -\mathbf{D}_{r_i}[\mathbf{r}_{-i}^{(t)}, \mathbf{r}_{-i}^{(-t)}] \ - \mathbf{e}_{r_i}[\mathbf{r}_{-i}^{(t)}] + \mathbf{h}_{r_i}[\mathbf{r}_{-i}^{(-t)}]$$

gives

$$p_1^{(t)}(\mathbf{s}) = \text{Sign}(\Delta[s_1^{(t)}, s_2^{(t)}]) - \text{Sign}(\Delta[s_1^{(t)} s_2^{(-t)}]),$$

which is just the payment rule of the CA mechanism.

4.1.3 The necessity of the condition. If a problem instance $\langle \mathbf{f}, M \rangle$ does not satisfy the condition in our theorem, then one should not hope for a strictly truthful scoring mechanism. The designer should seek additional knowledge about the distribution or elicit different information. Testing whether there exist power diagrams as described in Theorem 4.4 may not be easy. Below we provide two simpler necessary conditions for the existence of strictly truthful scoring mechanisms. When the report space $\mathcal{R}_i$ is discrete, these two necessary conditions can be implied by Theorem 4.4. In addition, they will also hold for continuous report space, for example, when $\mathcal{R}_i = \Delta\Omega$ is the space of all posteriors of the participant.

First, observe that a cell of a power diagram must be convex. Therefore, if we fix a marginal distribution $\mu(\mathbf{r}_{-i})$, $Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))$ for different $\mathbf{r}_i$ should fall into disjoint convex sets. More specifically, we should have the following.

Proposition 4.5. *A multi-task problem $\langle \mathbf{f}, M\rangle$ is not elicitable if there exist $i \in [n]$ and a marginal distribution $\mu(\mathbf{r}_{-i})$ and $\mathbf{r}_i \neq \mathbf{r}_i' \in \mathcal{R}_i$, such that there exist $k, l \in \mathbb{Z}^+$ and $\mathbf{q}_1, \dots, \mathbf{q}_k \in Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i}))$ and $\beta_1, \dots, \beta_k \in [0,1]$ with $\beta_1 + \cdots + \beta_k = 1$ and $\mathbf{q}_1', \dots, \mathbf{q}_l' \in Q_i(\mathbf{r}_i', \mu(\mathbf{r}_{-i}))$ and $\beta_1', \dots, \beta_l' \in [0,1]$ with $\beta_1' + \cdots + \beta_l' = 1$ such that*

$$\beta_1 \cdot \mathbf{q}_1 + \cdots + \beta_k \cdot \mathbf{q}_k = \beta_1' \cdot \mathbf{q}_1' + \cdots + \beta_l' \cdot \mathbf{q}_l'.$$

The second necessary condition is even simpler. It is similar to the robust stochastic relevance condition 3.3 that we proposed for the single-task peer prediction. The difference in the multi-task setting is that we need to fix a marginal distribution $\mu(\mathbf{r}_{-i})$.

Proposition 4.6 (Robust stochastic relevance with given marginal distributions). *For a multi-task problem $\langle \mathbf{f}, M\rangle$, if there exist $i \in [n]$ and a marginal distribution $\mu(\mathbf{r}_{-i})$ such that*

$$Q_i(\mathbf{r}_i, \mu(\mathbf{r}_{-i})) \cap Q_i(\mathbf{r}_i', \mu(\mathbf{r}_{-i})) \neq \emptyset$$

for $\mathbf{r}_i \neq \mathbf{r}_i' \in \mathcal{R}_i$, then $\langle \mathbf{f}, M\rangle$ is not elicitable.

Although Proposition 4.6 seems apparent based on Theorem 4.4, it can be useful when we consider a complicated report space, e.g., when we consider eliciting linear properties of the participants' posteriors (Section 5). It is the main tool that we use to prove our main results in Section 5.

4.2 General mechanisms

Although all the existing strictly truthful mechanisms that we know belong to scoring mechanisms, recently there have been attempts to design more complicated mechanisms with good properties, e.g., the mechanism proposed by Kong [8] is not a scoring mechanism. In this section, we look at general multi-task mechanisms and give a necessary condition for a multi-task problem to be elicitable, assuming that the participants use consistent strategies, which is a commonly used assumption by the multi-task peer prediction literature.

The necessary condition basically says that, given a joint distribution of the participants' reports without naming a participant i's report, the designer should at least be able to label participant i's report based on the distribution. For example, consider two participants who are asked to report a high signal or a low signal, i.e., $\mathcal{R}_1 = \mathcal{R}_2 = \{\text{high}, \text{low}\}$. Then given a possible joint distribution without the labels of the first participant's report, e.g. the distribution represented by the table in the left. The designer should be able to tell which of r_1 and r_1' is the high report.

	high	low
r_1	0.8	0.2
r_1'	0.2	0.8

$\Longrightarrow$

	high	low
high	0.8	0.2
low	0.2	0.8

Formally, we have the following theorem.

Theorem 4.7. *A multi-task problem $\langle \mathbf{f}, M\rangle$ is elicitable under consistent strategies only if for any distribution $\mu \in M$, any agent i and any permutation π of $\mathcal{R}_i$, there exists no $\widetilde{\mu} \in M$ such that*

$$\mu(\pi(r_i), \mathbf{r}_{-i}) = \widetilde{\mu}(r_i, \mathbf{r}_{-i}).$$

Or equivalently, for any i, if we use a $|\mathcal{R}_i| \times |\mathcal{R}_{-i}|$ matrix A_μ with $\mu(r_i, \mathbf{r}_{-i})$ in row r_i and column $\mathbf{r}_{-i}$ to represent any $\mu \in M$, then for any $\mu \in M$ and any $|\mathcal{R}_i| \times |\mathcal{R}_i|$ permutation matrix P, there exists no $\widetilde{\mu} \in M$ such that

$$P \cdot A_\mu = A_{\widetilde{\mu}}.$$

Proof. We prove that if there exist $\mu, \widetilde{\mu} \in M$ and permutation π that have $\mu(\pi(r_i), \mathbf{r}_{-i}) = \widetilde{\mu}(r_i, \mathbf{r}_{-i})$, there exists no payment rule $p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)})$ that guarantees strict truthfulness under consistent strategies for both μ and $\widetilde{\mu}$.

Since $\mu(\pi(r_i), \mathbf{r}_{-i}) = \widetilde{\mu}(r_i, \mathbf{r}_{-i})$, we have

$$\begin{aligned}
\mathbb{E}_{\widetilde{\mu}}[p_i(\pi(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)})] &= \sum_{\mathbf{r}^{(1:T)}} \widetilde{\mu}(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) \cdot p_i(\pi(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)}) \\
&= \sum_{\mathbf{r}^{(1:T)}} \mu(\pi(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)}) \cdot p_i(\pi(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)}) \\
&= \sum_{\mathbf{r}^{(1:T)}} \mu(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) \cdot p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) \\
&= \mathbb{E}_{\mu}[p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)})],
\end{aligned}$$

which means that the expected payment of permuting the reporting when the underlying distribution is $\widetilde{\mu}$ is equal to the expected payment of truthfully reporting when the distribution is μ. At the same time, it should also hold that

$$\begin{aligned}
\mathbb{E}_{\widetilde{\mu}}[p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)})] &= \sum_{\mathbf{r}^{(1:T)}} \widetilde{\mu}(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) \cdot p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) \\
&= \sum_{\mathbf{r}^{(1:T)}} \mu(\pi(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)}) \cdot p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) \\
&= \sum_{\mathbf{r}^{(1:T)}} \mu(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) \cdot p_i(\pi^{-1}(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)}) \\
&= \mathbb{E}_{\mu}[p_i(\pi^{-1}(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)})],
\end{aligned}$$

which means that the expected payment of truthfully reporting when the underlying distribution is $\widetilde{\mu}$ is equal to the expected payment of inversely permuting the strategy when the distribution is μ. Then it is easy to see that $p_i(\cdot)$ cannot be strictly truthful for both μ and $\widetilde{\mu}$. If $p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)})$ is strictly truthful for $\widetilde{\mu}$, then by definition, permuting the truthful reports should lead to strictly lower expected payment,

$$\mathbb{E}_{\widetilde{\mu}}[p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)})] > \mathbb{E}_{\widetilde{\mu}}[p_i(\pi(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)})],$$

which will violate the strict truthfulness for μ

$$\mathbb{E}_{\mu}[p_i(\pi^{-1}(r_i^{(1:T)}), \mathbf{r}_{-i}^{(1:T)})] > \mathbb{E}_{\mu}[p_i(r_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)})]$$

according to the two equalities above. □

One may wonder whether the permutation can be replaced by any reporting strategy, i.e., whether the permutation matrix can be replaced by any Markov matrix. The answer is negative. For the proof to hold, we need the matrix P to be invertible, and the only invertible Markov matrices are permutation matrices.

The previous works that design strictly truthful multi-task peer prediction mechanisms make assumptions that automatically satisfy this condition. To our knowledge, this necessary condition did not appear in any previous work, but similar techniques have been used in a different setting. Shnayder et al. [22] used a similar approach to prove that the CA mechanism is maximally strong truthful among a broader class of mechanisms (Theorem 5.9).

5 LINEAR PROPERTIES

In this section, we consider the reports that are linear in the participant's posterior $\mu(\omega|s_i)$. More specifically, the report function $\mathbf{r}_i$ is a length-L vector with

$$\mathbf{r}_i = \mathbf{G} \cdot \boldsymbol{\mu}(\omega|s_i)$$

where $\mathbf{G}$ is a $L \times |\Omega|$ matrix that represents the linear transformation from the posterior $\mu(\omega|s_i)$ to $\mathbf{r}_i$. Or equivalently, each entry of $\mathbf{r}_i$ is the expectation of a random variable defined on Ω. Common examples include

- the posterior itself $\mathbf{r}_i = \mu(\omega|s_i)$;
- the moments of the state ω when the state is a real number $\mathbb{E}[\omega|s_i], \mathbb{E}[\omega^2|s_i], \mathbb{E}[\omega^3|s_i] \dots$.

The elicitability of such linear properties has been studied by Abernethy and Frongillo [1] when the designer can observe ω and design a payment based on both the report and ω. They showed that linear properties are always elicitable in that case. But in peer prediction problems, the designer cannot observe ω but only have the participants' reports that are correlated with ω.

Our results show that it may not always be possible to elicit reports that are linear in the participants' posterior. We first consider two participants with signals independent conditioning on ω. In this case, if the designer has a single-task problem or the designer only considers scoring mechanisms for a multi-task problem, then the designer basically can only elicit linear properties that are equivalent to $\mu(\omega|s_i)$, assuming that the designer is uncertain about the underlying distribution. We then look into the case when the reports r_i is just the participants' posteriors $\mu(\omega|s_i)$. We give a necessary condition for $\mu(\omega|s_i)$ to be elicitable. This condition implies that the mechanisms proposed by [11] are already the best we can hope for in their setting, in the sense that their mechanisms can work for any problem instance that can possibly be elicitable.

5.1 Impossibility result for two agents

Suppose we have two participants whose signals are conditionally independent, that is, the joint distribution $\mu(\omega, s_1, s_2) = \mu(\omega)\mu(s_1|\omega)\mu(s_2|\omega)$ for all $\omega \in \Omega, s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2$. We assume that the designer is uncertain about the conditional distributions of the participants' signals $\mu(s_1|\omega), \mu(s_2|\omega)$. We formally define the uncertainty as follows.

Definition 5.1. *We say that a designer with an elicitability problem $\langle \mathbf{f}, M \rangle$ is minimally uncertain about the conditional distributions of the participants' signals if there exist a prior $\mu(\omega)$ and a set of possible conditional distributions $\mu(s_i|\omega)$ for each participant i, denoted by $M_i \subseteq \mathbb{R}^{|\mathcal{S}_i| \times |\Omega|}$, such that for any $\mu'(s_1|\omega) \in M_1, \dots, \mu'(s_n|\omega) \in M_n$, the joint distribution generated by them and the prior $\mu(\omega)$*

$$\mu'(\omega, s_1, \dots, s_n) = \mu(\omega)\mu'(s_1|\omega) \cdots \mu'(s_n|\omega), \ \forall \omega, s_1, \dots, s_n$$

is a possible joint distribution $\mu'(\omega, s_1, \dots, s_n) \in M$, and each M_i is open relative to the space of valid conditional distributions $(\Delta\mathcal{S}_i)^{|\Omega|}$.

The main restriction here is that M_i needs to be an open set relative to the space of valid conditional distributions. This means that for each participant i, there should exist a possible conditional distribution $\mu(s_i|\omega)$ such that the nearby conditional distributions are also possible.

Then if the designer never asks anyone for trivial reports that is a constant function of the participant's posterior, the only elicitable linear properties are the ones that are equivalent to the posteriors $\mu(\omega|s_i)$, assuming that the designer only wants scoring mechanisms for multi-task problems. More specifically, we have the following theorem.

Theorem 5.2. *For a single-task/multiple-task elicitability problem $\langle \mathbf{f}, M \rangle$ with $n = 2$ with conditional independent signals, if the designer is minimally uncertain about the conditional distributions*

of the participants' signals, and for any participant i, the truthful report $\mathbf{r}_i = \mathbf{G} \cdot \mu(\omega|s_i)$ is a non-constant linear function of the posterior $\mu(\omega|s_i)$, then $\mathbf{r}$ is elicitable/elicitable by scoring mechanisms only if there is a one-to-one mapping from $\mu(\omega|s_i)$ to $\mathbf{r}_i$ for all i, i.e.,

$$matrix \begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix} \text{ has rank } |\Omega|, \tag{10}$$

where $\mathbf{1}$ is the all-ones vector with length $|\Omega|$.

Here condition (10) is equivalent to that there exists a one-to-one mapping from $\mu(\omega|s_i)$ to $\mathbf{r}_i$ because if the matrix in (10) has rank $|\Omega|$, then it has linearly independent columns, which means that it has a left inverse H such that $H \cdot \begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix} = I$. Then we can recover $\mu(\omega|s_i)$ from $\mathbf{r}_i$ as $H \cdot \begin{bmatrix} \mathbf{r}_i \\ 1 \end{bmatrix} = H \cdot \begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix} \cdot \mu(\omega|s_i) = \mu(\omega|s_i)$.

According to Theorem 5.2, a designer who is interested in some information that is linear in the participants' posteriors should just try to elicit their posteriors $\mu(\omega|s_i)$, if the conditions in the theorem are satisfied. The proof of the theorem is based on the necessary condition in Corollary 4.6. The proof of Theorem 5.2 is quite involved. The high level idea is that if the columns of the matrix $\begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix}$ are linearly dependent, we can find two possible distribution $\mu, \mu' \in M$ so that there exist two signal realizations s_i, s_i' of a participant i that will lead to the same posterior in μ and μ'

$$\mu(\mathbf{r}_{-i}|s_i) = \mu'(\mathbf{r}_{-i}|s_i'),$$

but the truthful reports are different

$$\mathbf{G} \cdot \mu(\omega|s_i) \neq \mathbf{G} \cdot \mu'(\omega|s_i').$$

So it violates Proposition 4.6. We defer the full proof to Appendix D.1.

5.2 Prediction on the State

In this section, we study the report function that is just the participants' posteriors about ω after observing their signals $\mathbf{r}_i = \mu(\omega|s_i)$. We provide a necessary condition for $\mu(\omega|s_i)$ to be elicitable when the participants' signals are conditionally independent and the designer is uncertain about the conditional distributions of the participants' signals.

For each participant i, our necessary condition will look at the conditional distribution of the other participants' signals $\mu(\mathbf{s}_{-i}|\omega)$. For every $\mu \in M$, define $\mathbf{P}_i^\mu$ to be a $|\mathcal{S}_{-i}| \times |\Omega|$ matrix with entry in row $\mathbf{s}_{-i}$ and column ω equal to $\mu(\mathbf{s}_{-i}|\omega)$.

Theorem 5.3. *For a single-task/multiple-task elicitability problem $\langle \mathbf{f}, M \rangle$ with $\mathbf{r}_i = \mu(\omega|s_i)$ with conditional independent signals, if the designer is minimally uncertain about the conditional distributions of the participants' signals, then the condition*

$$rank(\mathbf{P}_i^\mu) = |\Omega|, \ \forall i, \forall \mu$$

is necessary for $\langle \mathbf{f}, M \rangle$ to be elicitable/elicitable by scoring mechanisms.

The proof can be found in Appendix D.2.

This necessary condition basically means that to guarantee that truthfully reporting is a strictly optimal strategy for participant i (at the equilibrium), the other participants' signals need to be sufficiently correlated with the state ω. Since Kong and Schoenebeck [11] proposed a mechanism that elicits $\mu(\omega|s_i)$ when the designer knows the prior $\mu(\omega)$ (see details in Section 3.3), the condition in Theorem 5.3 is also sufficient when the prior $\mu(\omega)$ is known.

Corollary 5.4. *For a single-task/multiple-task elicitability problem $\langle \mathbf{f}, M \rangle$ with $\mathbf{r}_i = \mu(\omega|s_i)$ with known prior $\mu(\omega)$ with conditionally independent signals, if the designer is minimally uncertain about the conditional distributions of the participants' signals, then the condition that $rank(\mathbf{P}_i^\mu) = |\Omega|$ for all i and μ is both necessary and sufficient for $\mathbf{r}_i = \mu(\omega|s_i)$ to be elicitable/elicitable by scoring mechanisms.*

The necessity of the condition implies that the mechanisms proposed by Kong and Schoenebeck [11] are able to solve all the problem instances that are elicitable (by scoring mechanisms in the multi-task setting) if we consider the problem of eliciting $\mu(\omega|s_i)$ when the prior $\mu(\omega)$ is known.

Corollary 5.5. *For a single-task/multiple-task elicitability problem $\langle \mathbf{f}, M \rangle$ with $\mathbf{r}_i = \mu(\omega|s_i)$ with known prior $\mu(\omega)$ with conditionally independent signals, if the designer is minimally uncertain about the conditional distributions of the participants' signals, then all the problem instances $\langle \mathbf{f}, M \rangle$ that are elicitable/elicitable by scoring mechanisms can be solved by the mechanism proposed by Kong and Schoenebeck [11], i.e., their mechanism will be strictly truthful for all $\mu \in M$.*

6 DISCUSSION

We study the elicitability of multi-task peer prediction problems. Our main contribution includes (1) we characterize the elicitable multi-task peer prediction problems when the designer only uses scoring mechanisms, (2) we are the first to study the elicitability of properties that are linear in the participants' posteriors. We believe that the most intriguing future direction is to further simplify our characterization and find more applications of this result, by either considering more specific settings or adopting more advanced tools. For example, our characterization does not impose any restriction on the designer's knowledge about the distribution: the set M can be an arbitrary set of possible distributions. An immediate question is: can we simplify the characterization if M has a certain structure? Our negative result for linear properties only used a simplified version of the necessary condition. We believe that stronger results can be proved if we deploy more of the structure of power diagrams, e.g. the convexity of the cells. Finally, our result shows that it is possible to have a simpler characterization by restricting the class of mechanisms. It may be possible to simplify our result by considering other classes of mechanisms.

ACKNOWLEDGMENTS

The authors would like to thank all the anonymous reviewers for their careful reading, valuable comments, and constructive remarks. This work is supported by the National Science Foundation under Grant No. IIS 2007887.

REFERENCES

[1] Jacob D. Abernethy and Rafael M. Frongillo. 2012. A Characterization of Scoring Rules for Linear Properties. In *Proceedings of the 25th Annual Conference on Learning Theory*, Vol. 23. 27.1–27.13.

[2] Yiling Chen, Yiheng Shen, and Shuran Zheng. 2020. Truthful Data Acquisition via Peer Prediction. arXiv:cs.GT/2006.03992

[3] Anirban Dasgupta and Arpita Ghosh. 2013. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd international conference on World Wide Web*. 319–330.

[4] Boi Faltings and Goran Radanovic. 2017. Game theory for data science: Eliciting truthful information. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 11, 2 (2017), 1–151.

[5] Rafael Frongillo and Jens Witkowski. 2017. A geometric perspective on minimal peer prediction. *ACM Transactions on Economics and Computation (TEAC)* 5, 3 (2017), 1–27.

[6] Rafael M Frongillo. 2013. *Eliciting private information from selfish agents*. Ph.D. Dissertation. UC Berkeley.

[7] Radu Jurca and Boi Faltings. 2008. Incentives for expressing opinions in online polls. In *Proceedings of the 9th ACM Conference on Electronic Commerce*. 119–128.

[8] Yuqing Kong. 2020. Dominantly truthful multi-task peer prediction with a constant number of tasks. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2398–2411.

[9] Yuqing Kong. 2021. Counting the Number of People That Are Less Clever Than You. *arXiv preprint arXiv:2103.02214* (2021).

[10] Yuqing Kong and Grant Schoenebeck. 2018. Equilibrium Selection in Information Elicitation without Verification via Information Monotonicity. In *9th Innovations in Theoretical Computer Science Conference*.

[11] Yuqing Kong and Grant Schoenebeck. 2018. Water from two rocks: Maximizing the mutual information. In *Proceedings of the 2018 ACM Conference on Economics and Computation*. 177–194.

[12] Yuqing Kong and Grant Schoenebeck. 2019. An information theoretic framework for designing information elicitation mechanisms that reward truth-telling. *ACM Transactions on Economics and Computation (TEAC)* 7, 1 (2019), 2.

[13] Nicolas Lambert and Yoav Shoham. 2009. Eliciting truthful answers to multiple-choice questions. In *Proceedings of the 10th ACM conference on Electronic commerce*. 109–118.

[14] Nicolas S. Lambert, David M. Pennock, and Yoav Shoham. 2008. Eliciting Properties of Probability Distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce (EC '08)*. Association for Computing Machinery, New York, NY, USA, 129–138. https://doi.org/10.1145/1386790.1386813

[15] Yang Liu, Juntao Wang, and Yiling Chen. 2020. Surrogate scoring rules. In *Proceedings of the 21st ACM Conference on Economics and Computation*. 853–871.

[16] N. Miller, P. Resnick, and R. Zeckhauser. 2005. Eliciting informative feedback: The peer-prediction method. *Management Science* (2005), 1359–1373.

[17] D. Prelec. 2004. A Bayesian Truth Serum for subjective data. *Science* 306, 5695 (2004), 462–466.

[18] Goran Radanovic and Boi Faltings. 2013. A robust bayesian truth serum for non-binary signals. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI" 13)*. 833–839.

[19] Goran Radanovic and Boi Faltings. 2014. Incentives for truthful information elicitation of continuous signals. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI" 14)*. 770–776.

[20] Grant Schoenebeck and Fang-Yi Yu. 2020. Learning and Strongly Truthful Multi-Task Peer Prediction: A Variational Approach. arXiv:cs.GT/2009.14730

[21] Grant Schoenebeck and Fang-Yi Yu. 2020. Two Strongly Truthful Mechanisms for Three Heterogeneous Agents Answering One Question. In *International Conference on Web and Internet Economics*. Springer.

[22] Victor Shnayder, Arpit Agarwal, Rafael Frongillo, and David C Parkes. 2016. Informed truthfulness in multi-task peer prediction. In *Proceedings of the 2016 ACM Conference on Economics and Computation*. 179–196.

[23] J. Witkowski and D. Parkes. 2013. Learning the Prior in Minimal Peer Prediction.

[24] Jens Witkowski and David C. Parkes. 2012. Peer prediction without a common prior. In *Proceedings of the 13th ACM Conference on Electronic Commerce, EC 2012, Valencia, Spain, June 4-8, 2012*, Boi Faltings, Kevin Leyton-Brown, and Panos Ipeirotis (Eds.). ACM, 964–981. https://doi.org/10.1145/2229012.2229085

[25] Peter Zhang and Yiling Chen. 2014. Elicitability and knowledge-free elicitation with peer prediction. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. 245–252.

A TABLE OF NOTATIONS

symbols	single tasks	multiple tasks
space of signal	$\boldsymbol{\mathcal{S}} := \times_i \mathcal{S}_i$, and $\mathbf{s} \in \boldsymbol{\mathcal{S}}$	$\boldsymbol{\mathcal{S}} := \times_i \mathcal{S}_i$, and $\mathbf{s}^{(1:T)} \in \boldsymbol{\mathcal{S}}^T$
common prior	μ a distribution on $\Omega \times \mathcal{S}_1 \cdots \times \mathcal{S}_n$	$\mu^{\otimes T}$, each task is identically and independently sampled from μ
Designer's knowledge	M a subset of distributions that contains the underlying distribution $\mu \in M$	same
report space	$\boldsymbol{\mathcal{R}} := \times_i \mathcal{R}_i$	same
private signal	$\mathbf{S} := (S_i)_i$ the random variable of signals sampled from μ	$\mathbf{S}^{(1:T)} := (\mathbf{S}_i^{(1:T)})_i = (S_i^{(t)})_{i,t}$ a random variable sampled from $\mu^{\otimes T}$
truthful map	$\mathbf{f} = (f_{i,\mu})_{i,\mu}$	same
truthful report	$\mathbf{R} = (R_i)_i$, a random variable $(f_{i,\mu}(S_i))_i$ where $(S_i)_i \sim \mu$	$\mathbf{R}^{(1:T)} = (R_i^{(1:T)})_i = (R_i^{(t)})_{i,t}$, a random variable
strategy	$\boldsymbol{\sigma} = (\sigma_i)_i$, a collection of n independent random mappings where $\sigma_i : \mathcal{S}_i \to \Delta(\mathcal{R}_i)$.	$\boldsymbol{\sigma} = (\sigma_i^{(1:T)})_i = (\sigma_i^{(t)})_{i,t}$ a collection of nT independent random mapping $\sigma_i^{(t)} : \mathcal{S}_i \to \Delta(\mathcal{R}_i)$
strategic report	$\widetilde{\mathbf{R}} = (\widetilde{R}_i)_i = (\sigma_i(S_i))_i$, random variables $(\sigma_i(S_i))_i$ depends on $(\sigma_i)_i$ and $(S_i)_i$	$\widetilde{\mathbf{R}}^{(1:T)} = (\widetilde{R}_i^{(1:T)})_i = (\widetilde{R}_i^{(t)})_{i,t} = (\sigma_i^{(t)}(S_i^{(t)}))_{i,t}$
payment	$\mathbf{p} = (p_i)_i$ where $p_i : \times_i \mathcal{R}_i \to \mathbb{R}$	$\mathbf{p} = (p_i)_i$ where $p_i : \times_i \mathcal{R}_i^T \to \mathbb{R}$

B MISSING PROOFS IN SECTION 3

Given a distribution λ on a finite set $\Omega \times \mathcal{S}_1 \times \cdots \times \mathcal{S}_n$ and $\epsilon > 0$, let $B(\lambda, \epsilon) = \{\mu : \|\mu - \lambda\|_1 \le \epsilon\} \subset \Delta(\Omega \times \mathcal{S}_1 \times \cdots \times \mathcal{S}_n)\}$ be the collection of distributions whose total variation distance from λ is smaller than ϵ. Note that if $\epsilon = 0$, $B(\lambda, 0)$ is a singleton and we know exactly the true distribution. There are several truthful mechanisms, e.g., Miller et al. [16]. The following theorem shows we can still have truthful mechanisms when $\epsilon > 0$ is small enough.

Theorem B.1. *If $f_i(s_i) = s_i$ for all i and $\lambda \in \Delta(\Omega \times \mathcal{S}_1, \ldots \mathcal{S}_n)$ has full support $\lambda > 0$ and is stochastic relevant so that $\lambda(\mathbf{s}_{-i}|s_i) \neq \lambda(\mathbf{s}_{-i}|s_i')$ for all $i \in [n]$ and distinct s_i and s_i', there exists $\epsilon > 0$ such that $\langle \mathbf{f}, B(\lambda, \epsilon) \rangle$ is elicitable.*

The idea is very similar to the maximal robust mechanisms in Frongillo and Witkowski [5]. However, instead of the joint distribution being close to a center λ, their result requires all conditional distributions are close to a center. The proof is straightforward and it is in the appendix.

Proof of Corollary 3.3. If $Q_i(r_i) \cap Q_i(r_i') \neq \emptyset$, then there exist μ_1 and μ_2 with

$$\mu_1(\mathbf{r}_{-i}|s_i) = \mu_2(\mathbf{r}_{-i}|s_i'), \quad f(s_i) = r_i, \quad f(s_i') = r_i'.$$

Then if there exists a truthful payment rule $\mathbf{p}(\mathbf{r})$, then by the definition of elicitability Definition 2.7 it should satisfy that

$$\mathbb{E}_{\mathbf{r}_{-i} \sim \mu_1(\mathbf{r}_{-i}|s_i)}[p_i(r_i, \mathbf{r}_{-i})] > \mathbb{E}_{\mathbf{r}_{-i} \sim \mu_1(\mathbf{r}_{-i}|s_i)}[p_i(r_i', \mathbf{r}_{-i})],$$

$$\mathbb{E}_{\mathbf{r}_{-i}\sim\mu_2(\mathbf{r}_{-i}|s_i')}[p_i(r_i',\mathbf{r}_{-i})] > \mathbb{E}_{\mathbf{r}_{-i}\sim\mu_2(\mathbf{r}_{-i}|s_i')}[p_i(r_i,\mathbf{r}_{-i})].$$

which is impossible if $\mu_1(\mathbf{r}_{-i}|s_i) = \mu_2(\mathbf{r}_{-i}|s_i')$. □

Proof of Theorem B.1. Because λ is stochastic relevant, $\langle \mathbf{f}, \{\lambda\}\rangle$ is elicitable. [16]

By Theorem 3.2, for agent i, there exists a power diagram with sites $\mathbf{v}^r \in \mathbb{R}^{|\mathcal{R}_{-i}|}$ and weights $w^r \in \mathbb{R}$ for all $r \in \mathcal{R}_i$ such that for all $r_i \in \mathcal{R}_i$, $r_i = \arg\min_s\{\langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^s\rangle - w^s\}$, and for any distinct pair r_i, r_i'

$$\langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i}\rangle - w^{r_i} > \langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i'}\rangle - w^{r_i'}$$

Now we want to prove the same power diagram works for $\langle \mathbf{f}, B(\lambda,\epsilon)\rangle$ when ϵ is small enough: For all $r_i \in \mathcal{R}_i$ and $\mu \in B(\lambda,\epsilon)$, $cell(\mathbf{v}^{r_i})$ contains $Q_i(r_i)$.

First we can bound the distance between conditional distributions by the distance between their joint distributions. For any $\mu > 0$ and $r_i \in \mathcal{R}_i$

$$\begin{aligned}
\|\mu(\mathbf{r}_{-i}|r_i) - \lambda(\mathbf{r}_{-i}|r_i)\|_1 &= \sum_{\mathbf{r}_{-i}}\left|\frac{\mu(r_i,\mathbf{r}_{-i})}{\mu(r_i)} - \frac{\lambda(r_i,\mathbf{r}_{-i})}{\lambda(r_i)}\right| \\
&\le \sum_{\mathbf{r}_{-i}}\left|\frac{\mu(r_i,\mathbf{r}_{-i})}{\mu(r_i)} - \frac{\mu(r_i,\mathbf{r}_{-i})}{\lambda(r_i)}\right| + \sum_{\mathbf{r}_{-i}}\left|\frac{\mu(r_i,\mathbf{r}_{-i})}{\lambda(r_i)} - \frac{\lambda(r_i,\mathbf{r}_{-i})}{\lambda(r_i)}\right| \\
&\le |\mu(r_i)^{-1} - \lambda(r_i)^{-1}|\sum_{\mathbf{r}_{-i}}|\mu(r_i,\mathbf{r}_{-i})| + \lambda(r_i)^{-1}\sum_{\mathbf{r}_{-i}}|\mu(r_i,\mathbf{r}_{-i}) - \lambda(r_i,\mathbf{r}_{-i})| \\
&\le |\mu(r_i)^{-1} - \lambda(r_i)^{-1}| + \lambda(r_i)^{-1}\|\mu-\lambda\|_1
\end{aligned}$$

Now it is sufficient to bound the first term by $\|\mu-\lambda\|_1$. If $\|\mu-\lambda\|_1 < \epsilon$, we can take ϵ smaller than $\min_{r_i}\lambda(r_i)/2$. Thus, for all r_i, we have $|\mu(r_i) - \lambda(r_i)| < \lambda(r_i)/2$, and $|\mu(r_i)^{-1} - \lambda(r_i)^{-1}| = |\mu(r_i) - \lambda(r_i)|/(\mu(r_i)\lambda(r_i)) \le \frac{2}{\lambda(r_i)^2}\|\mu-\lambda\|_1$. Therefore, we have

$$\|\mu(\mathbf{r}_{-i}|r_i) - \lambda(\mathbf{r}_{-i}|r_i)\|_1 \le \left(\frac{2}{\lambda(r_i)^2} + \frac{1}{\lambda(r_i)}\right)\|\mu-\lambda\|_1 \le \frac{3}{\lambda(r_i)^2}\|\mu-\lambda\|_1 \tag{11}$$

On the other hand, since $\mathcal{R}_i$ is a finite set and $\lambda > 0$, we can pick $\epsilon_0 > 0$ small enough such that for all $r_i \ne r_i'$

$$3\epsilon_0 + \langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i}\rangle - w^{r_i} < \langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i'}\rangle - w^{r_i'}. \tag{12}$$

Now for all distinct r_i, r_i' and $\mu \in B(\lambda, \min_{r_i}\frac{\lambda(r_i)^2}{\|\mathbf{v}^{r_i}\|_\infty}\epsilon_0)$, we have

$$\begin{aligned}
&\left(\langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i'}\rangle - w^{r_i'}\right) - \left(\langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i}\rangle - w^{r_i}\right) \\
=&\langle\lambda(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i'} - \mathbf{v}^{r_i}\rangle - w^{r_i'} + w^{r_i} \\
\ge&\langle\lambda(\mathbf{r}_{-i}|r_i) - \mu(\mathbf{r}_{-i}|r_i), \mathbf{v}^{r_i'} - \mathbf{v}^{r_i}\rangle + 3\epsilon_0 && \text{(by Eqn. (12))} \\
\ge&3\epsilon_0 - \|\lambda(\mathbf{r}_{-i}|r_i) - \mu(\mathbf{r}_{-i}|r_i)\|_1\|\mathbf{v}^{r_i}\|_\infty && \text{(Holder's inequality)} \\
>&3\epsilon_0 - \frac{3\|\mathbf{v}^{r_i}\|_\infty}{\lambda(r_i)^2}\|\lambda-\mu\|_1 && \text{(by Eqn. (11))} \\
\ge&0 && \left(\|\mu-\lambda\|_1 \le \tfrac{\lambda(r_i)^2}{\|\mathbf{v}^{r_i}\|_\infty}\epsilon_0\right)
\end{aligned}$$

This completes our proof. □

C MISSING PROOFS IN SECTION 4

C.1 Proof of Theorem 4.1

First it is apparent that a mechanism that guarantees strict truthfulness also guarantees strict truthfulness under consistent strategies. In the following proposition we show if there is a scoring mechanism so that a multi-task problem $\langle \mathbf{f}, M \rangle$ is elicitable under consistent strategies there is a scoring mechanism such that $\langle \mathbf{f}, M \rangle$ is elicitable under general strategies defined in Definition 2.3.

Proposition C.1. *If there exists a scoring rule mechanism* $\mathbf{p} = \{p_i^t : i \in [n], t \in [T]\}$ *so that a multi-task problem* $\langle \mathbf{f}, M \rangle$ *is elicitable under consistent strategies, there is a scoring mechanism such that* $\langle \mathbf{f}, M \rangle$ *is elicitable under general strategies.*

Proof of Proposition C.1. We construct such scoring mechanism against general strategies through symmetrization.

Since $\mathbf{p}$ is a scoring mechanism, agent i's payment p_i can be decomposed as $p_i(\mathbf{r}_i^{(1:T)}, \mathbf{r}_{-i}^{(1:T)}) = \sum_{t=1}^{T} p_i^{(t)}(r_i^{(t)}, \mathbf{r}_{-i}^{(1:T)})$ for all $\mathbf{r} \in \mathcal{R}$ where for agent i, $p_i^{(t)}$ only depends on his report on task i. We define a new payment $\hat{p}_i$ through symmetrization: Let $\mathrm{Sym}(T)$ be the collection of all permutations on set $[T]$.

$$\hat{p}_i(\mathbf{r}) := \frac{1}{T!} \sum_{\tau \in [T], \pi \in \mathrm{Sym}(T)} p_i^{(\tau)} \left(r_i^{(\pi(1))}, r_{-i}^{(\pi(1))}, r_{-i}^{(\pi(2))}, \ldots, r_{-i}^{(\pi(T))} \right).$$

First it is easy to see if all agents use consistent strategies $\boldsymbol{\sigma}$, the expectation of $\hat{p}_i$ equals the expectation of p_i

$$\mathbb{E}_{\boldsymbol{\sigma}, \mathrm{S}} \left[\hat{p}_i(\widetilde{\mathbf{R}}) \right] = \mathbb{E}_{\boldsymbol{\sigma}, \mathrm{S}} \left[p_i(\widetilde{\mathbf{R}}) \right]. \tag{13}$$

Formally, because when $\boldsymbol{\sigma}$ is consistent the distribution of $\widetilde{\mathbf{R}}$ is exchangeable on tasks so that the distribution of $(\widetilde{R}^{(1)} \ldots \widetilde{R}^{(T)})$ is identical to the distribution of $(\widetilde{R}^{(\pi(1))} \ldots \widetilde{R}^{(\pi(T))})$ for any permutation $\pi \in \mathrm{Sym}(T)$, the expectation of $\hat{p}_i$ is the sum of $T!$ identical terms, and Eqn. (13) holds.

Then we show that given any agent i's general strategy $\sigma_i = (\sigma_i^1, \ldots, \sigma_i^T)$, if all other agents are truth telling, there exists a consistent

$$\hat{\sigma}_i := \frac{1}{T} \sum_t \sigma_i^t$$

such that agent i's expected payment of $\hat{p}_i$ under strategy σ_i is equal to the expected payment of p_i under $\hat{\sigma}_i$,

$$\mathbb{E}_{\sigma_i, \mathrm{S}} \left[\hat{p}_i \left(\widetilde{\mathbf{R}}^{(1:T)}, \mathbf{R}_{-i}^{(1:T)} \right) \right] = \mathbb{E}_{\hat{\sigma}_i, \mathrm{S}} \left[p_i \left(\widetilde{\mathbf{R}}^{(1:T)}, \mathbf{R}_{-i}^{(1:T)} \right) \right] \tag{14}$$

Note that if Eqn. (14) holds, with Eqn. (13) we completes the proof. When every other agents are truth telling, the expected payment under any general nontruthful strategy σ_i is strictly less than the expected payment of truth telling,

$$\begin{aligned}
&\mathbb{E}_{\sigma_i, \mathrm{S}} \left[\hat{p}_i \left(\widetilde{\mathbf{R}}^{(1:T)}, \mathbf{R}_{-i}^{(1:T)} \right) \right] \\
=&\mathbb{E}_{\hat{\sigma}_i, \mathrm{S}} \left[p_i \left(\widetilde{\mathbf{R}}^{(1:T)}, \mathbf{R}_{-i}^{(1:T)} \right) \right] && \text{(by (14))} \\
<&\mathbb{E}_{\mathrm{S}} \left[p_i \left(\widetilde{\mathbf{R}} \right) \right] && (p_i \text{ is truthful under consistent strategies.}) \\
=&\mathbb{E}_{\mathrm{S}} \left[\hat{p}_i \left(\widetilde{\mathbf{R}} \right) \right] && \text{(by (13) and truth telling is an consistent strategy)}
\end{aligned}$$

The above inequality is strict, because if σ_i is not truthful, the average $\hat{\sigma}_i$ is also not truthful.

Finally, let's prove Eqn. (14). We set $\mathbf{R}_{-i}^{\otimes(T-1)}$ be a sequence of $T-1$ iid truthful report on a generic tasks, and $(\mathbf{R}_{-i}^{(\pi(2))}, \ldots, \mathbf{R}_{-i}^{(\pi(T))})$ has the same distribution as $\mathbf{R}_{-i}^{\otimes(T-1)}$ for any permutation $\pi \in \mathrm{Sym}(T)$, because the distribution is exchangeable. With this notion, due to the linearity of expectation, we have

$$
\begin{aligned}
\mathbb{E}_{\sigma_i,\mathrm{S}}[\hat{p}_i(\widetilde{\mathbf{R}})] =& \frac{1}{T}\mathbb{E}\left[\sum_{\tau,\tau'} p_i^{(\tau)}\left(\sigma_i^{(\tau')}(S_i^{(\tau')}), \mathbf{R}_{-i}^{(\tau')}, \mathbf{R}_{-i}^{\otimes(T-1)}\right)\right] \\
=& \frac{1}{T}\mathbb{E}_{\sigma_i,\mathrm{S}}\left[\sum_{\tau,\tau'} p_i^{(\tau)}\left(\sigma_i^{\tau'}(S_i'), f_{-i,\mu}(S_{-i}'), \mathbf{R}_{-i}^{\otimes(T-1)}\right)\right] \quad \text{(Let } (S_i', S_{-i}') \text{ be sampled from } \mu.\text{)} \\
=& \frac{1}{T}\mathbb{E}_{\mathrm{S}}\left[\sum_{\tau,\tau'}\sum_{r_i\in\mathcal{R}_i} \Pr\left[\sigma_i^{(\tau')}(S_i') = r_i\right] p_i^{(\tau)}\left(r_i, f_{-i,\mu}(S_{-i}'), \mathbf{R}_{-i}^{\otimes(T-1)}\right)\right] \\
=& \mathbb{E}_{\mathrm{S}}\left[\sum_{\tau}\sum_{r_i\in\mathcal{R}_i}\left(\frac{1}{T}\sum_{\tau'}\Pr\left[\sigma_i^{(\tau')}(S_i') = r_i\right]\right) p_i^{(\tau)}\left(r_i, f_{-i,\mu}(S_{-i}'), \mathbf{R}_{-i}^{\otimes(T-1)}\right)\right] \\
=& \mathbb{E}_{\mathrm{S}}\left[\sum_{\tau}\sum_{r_i\in\mathcal{R}_i} \Pr\left[\hat{\sigma}_i(S_i') = r_i\right] p_i^{(\tau)}\left(r_i, f_{-i,\mu}(S_{-i}'), \mathbf{R}_{-i}^{\otimes(T-1)}\right)\right] \qquad (\hat{\sigma}_i = \tfrac{1}{T}\textstyle\sum_t \sigma_i^t) \\
=& \mathbb{E}_{\hat{\sigma}_i,\mathrm{S}}\left[\sum_{\tau} p_i^{(\tau)}\left(\hat{\sigma}_i(S_i'), f_{-i,\mu}(S_{-i}'), \mathbf{R}_{-i}^{\otimes(T-1)}\right)\right] = \mathbb{E}_{\hat{\sigma}_i,\mathrm{S}}\left[p_i(\widetilde{\mathbf{R}})\right]
\end{aligned}
$$

which completes the proof. □

C.2 Characterization for polynomial-size scoring mechanisms

Suppose the designer only uses scoring mechanisms with

$$p_i^{(t)}(\vec{r}_i^{(t)}, \vec{\mathbf{r}}_{-i}^{(1:T)}) = p_i^{(t)}(\vec{r}_i^{(t)}, \vec{r}_{-i}^{(t)}, Y(\vec{\mathbf{r}}_{-i}^{(-t)})),$$

where Y is an arbitrary function with range $\mathcal{Y}$. And $\mathcal{Y}$ has a polynomial size. We first define the power diagram constraint for fixed marginal distribution of $\mu(Y(\mathbf{r}_{-i}^{(1:T-1)}))$. The definition is similar to Definition 4.2. To simplify the notation, we write $Y(\mathbf{r}_{-i}^{(1:T-1)})$ as Y.

Define $M(\mu(Y))$ to be the set of all distributions that has marginal distribution of Y equal to $\mu(Y)$,

$$M(\mu(Y)) = \left\{\mu' \in M : \mu'(Y) = \mu(Y)\right\}$$

Also define $Q_i(r_i, \mu(Y))$ to be the set of participant i's possible posteriors about the others' reports when participant i's truthful report is r_i and the marginal distribution of Y is $\mu(Y)$,

$$Q_i(r_i, \mu(Y)) = \{\mu(\mathbf{r}_{-i}|s_i) : \mu \in M(\mu(Y)),\ s_i \text{ satisfies } f_{i,\mu}(s_i) = r_i\}.$$

Then the power diagram constraint for fixed marginal distribution of $\mu(Y)$ is defined as follows.

Definition C.2. *A multi-task problem $\langle \mathbf{f}, M\rangle$ satisfies the* power diagram constraint for fixed marginal distributions of $\mu(Y)$ *if for all i and $\mu(Y)$, $\{Q_i(r_i, \mu(Y))\}_{r_i\in\mathcal{R}_i}$ can be fitted into a power diagram, which means that there exists a power diagram in dimension $|\mathcal{R}_{-i}|$ with $|\mathcal{R}_i|$ cells defined by $\{\mathbf{v}^{r_i} \in \mathbb{R}^{|\mathcal{R}_{-i}|} : r_i \in \mathcal{R}_i\}$ with associated weights $\{w^{r_i} \in \mathbb{R} : r_i \in \mathcal{R}_i\}$, such that each $Q_i(r_i, \mu(Y))$ falls into a distinct cell,*

$$Q_i(r_i, \mu(Y)) \subseteq \mathit{cell}(\mathbf{v}^{r_i}),\ \text{for all } r_i \in \mathcal{R}_i.$$

Then we have the following theorem.

Theorem C.3. *If the designer only uses scoring mechanisms with*

$$p_i^{(t)}(\vec{r}_i^{(t)}, \vec{\mathbf{r}}_{-i}^{(1:T)}) = p_i^{(t)}(\vec{r}_i^{(t)}, \vec{r}_{-i}^{(t)}, Y(\vec{\mathbf{r}}_{-i}^{(-t)})),$$

where Y is an arbitrary function with range $\mathcal{Y}$, then a multi-task problem is elicitable if and only if (1) it satisfies the power diagram constraint for fixed marginal distribution of $\mu(Y)$; (2) the site $\mathbf{v}^{r_i}$ for different $\mu(Y)$ is an affine function of $\mu(Y)$ and the weights w^{r_i} for different $\mu(Y)$ is also an affine function of $\mu(Y)$. Formally, for every $i \in [n]$ and every $r_i \in \mathcal{R}_i$, there exist a matrix $\mathbf{D}_{r_i}$ with $|\mathcal{R}_{-i}|$ rows and $|\mathcal{Y}|$ columns and a vector $\mathbf{e}_{r_i} \in \mathbb{R}^{|\mathcal{R}_{-i}|}$ with

$$\mathbf{v}^{r_i}(\mu(Y)) = \mathbf{D}_{r_i} \cdot \mu(Y) + \mathbf{e}_{r_i},$$

and there exists a vector $\mathbf{h}_{r_i}$ with length $|\mathcal{Y}|$ such that

$$w^{r_i}(\mu(Y)) = \mathbf{h}_{r_i}^\top \cdot \mu(Y).$$

Moreover, if such power diagrams and affine functions exist, we can find a mechanism that is strictly truthful for any $\mu \in M$ with payments defined by the entries of $\mathbf{D}, \mathbf{e}$ and $\mathbf{h}$ as

$$p_i^{(t)}(r_i, \mathbf{r}_{-i}^{(t)}, Y(\vec{\mathbf{r}}_{-i}^{(-t)})) = -\mathbf{D}_{r_i}[\mathbf{r}_{-i}^{(t)}, Y(\vec{\mathbf{r}}_{-i}^{(-t)})] \ - \mathbf{e}_{r_i}[\mathbf{r}_{-i}^{(t)}] + \mathbf{h}_{r_i}[Y(\vec{\mathbf{r}}_{-i}^{(-t)})], \quad \forall r_i \in \mathcal{R}_i,\ \mathbf{r}_{-i}^{(t)},\ \mathbf{r}_{-i}^{(-t)}.$$

The proof of the theorem is entirely similar to the proof of Theorem 4.4.

D MISSING PROOFS IN SECTION 5

D.1 Proof of Theorem 5.2

Before starting the main proof, we give a useful lemma about distributions with conditional independent signals.

Lemma D.1. *If a distribution $\mu(\omega, s_1, s_2)$ has conditional independent signals, i.e., $\mu(\omega, s_1, s_2) = \mu(\omega)\mu(s_1|\omega)\mu(s_2|\omega)$, then participant 1's posterior about s_2 satisfies*

$$\mu(s_2|s_1) = \sum_{\omega \in \Omega} \mu(s_2|\omega)\mu(\omega|s_1).$$

Proof. It is because

$$\mu(s_2|s_1) = \sum_{\omega} \mu(s_2, \omega|s_1) = \sum_{\omega} \mu(s_2|\omega, s_1)\mu(\omega|s_1) = \sum_{\omega} \mu(s_2|\omega)\mu(\omega|s_1).$$

Here in the last equation, we've used the conditional independence: $\mu(s_2|\omega, s_1) = \frac{\mu(\omega, s_1, s_2)}{\mu(\omega, s_1)} = \frac{\mu(\omega)\mu(s_1|\omega)\mu(s_2|\omega)}{\mu(\omega)\mu(s_1|\omega)} = \mu(s_2|\omega)$. □

Consider a multi-task elicitability problem $\langle \mathbf{f}, M \rangle$ with two agents and with conditional independent signals with $\mathbf{r}_i = \mathbf{G} \cdot \boldsymbol{\mu}(\omega|s_i)$, and the designer is minimally uncertain about the conditional distributions of the participants' signals. We show that if $\text{rank}([\mathbf{G}^\top, \mathbf{1}]) < |\Omega|$, there exists a pair of distributions $\mu^*, \mu' \in M$ that will violate the condition in Proposition 4.6.

We first choose a distribution $\mu^* \in M$ as follows. Since the designer is minimally uncertain about the conditional distributions of the participants' signals, there exists a prior $\mu(\omega)$ and a set of possible conditional distributions for each participant $M_i \subseteq \mathbb{R}^{|\mathcal{S}_i| \times |\Omega|}$, such that the joint distributions generated by them are all possible, and each M_i is open relative to the space of valid conditional distributions $(\Delta \mathcal{S}_i)^{|\Omega|}$. Suppose that the first agent's set of possible likelihood functions $M_1 \subseteq \mathbb{R}^{|\mathcal{S}_1| \times |\Omega|}$ contains a ball with radius ε centered at $\mu^*(s_1|\omega)$, and M_2 contains a ball centered at $\mu^*(s_2|\omega)$.[3] Let $\mu^*(\omega, s_1, s_2)$ be the joint distribution generated by $\mu(\omega)$, $\mu^*(s_1|\omega)$ and $\mu^*(s_2|\omega)$. Then

[3]Here M_i contains a ball means that there is a ball $\mathcal{B} \subseteq \mathbb{R}^{|\mathcal{S}_i| \times |\Omega|}$ such that the intersection of the ball and the space of valid conditional distributions is still in M_i, i.e., $\mathcal{B} \cap (\Delta \mathcal{S}_i)^{|\Omega|} \subseteq M_i$.

for a certain $s_1^* \in \mathcal{S}_1$, participant 1's possible posterior about ω after seeing s_1^* must contain a ball centered at

$$\mu^*(\omega|s_1^*) \propto \mu(\omega)\mu^*(s_1^*|\omega)$$

with some radius $\varepsilon^* > 0$.

We then find another distribution $\mu' \in M$ such that

(a) the marginal distribution of $\mathbf{r}_2$ remains the same, $\mu'(\mathbf{r}_2) = \mu^*(\mathbf{r}_2)$;
(b) participant 1's truthful report after seeing s_1^* is different, $\mathbf{G} \cdot \boldsymbol{\mu}'(\omega|s_1^*) \neq \mathbf{G} \cdot \boldsymbol{\mu}^*(\omega|s_1^*)$;
(c) but participant 1's posterior about $\mathbf{r}_2$ after seeing s_1^* is unchanged, $\mu'(\mathbf{r}_2|s_1^*) = \mu^*(\mathbf{r}_2|s_1^*)$.

Then according to Proposition 4.6, $\mathbf{r}$ is not elicitable.

Since the report function $\mathbf{r}_i = \mathbf{G} \cdot \boldsymbol{\mu}(\omega|s_i)$ is nonconstant, there exists a length-$|\Omega|$ vector $\boldsymbol{\beta}$ with $\mathbf{G} \cdot \boldsymbol{\beta} \neq 0$ and $\mathbf{1}^\top \cdot \boldsymbol{\beta} = 0$. Since the possible posterior of the first agent after seeing s_1^* contains a ball centered at $\mu^*(\omega|s_1^*)$, there exists a small enough δ such that $\mu'(\omega|s_1^*) = \mu^*(\omega|s_1^*) + \delta \cdot \boldsymbol{\beta}$ is a possible posterior of the first agent and condition (b) is satisfied, i.e.,

$$\mathbf{G} \cdot \boldsymbol{\mu}^*(\omega|s_1^*) \neq \mathbf{G} \cdot \boldsymbol{\mu}'(\omega|s_1^*). \tag{15}$$

We will use this vector $\boldsymbol{\beta}$ to construct a likelihood function $\mu'(s_2|\omega) \in M_2$ so that (a) and (c) are also satisfied.

Let $\mathbf{Q}^*$ be an $|\mathcal{S}_2| \times |\Omega|$ matrix with $\mu^*(s_2|\omega)$ at row s_2 and column ω. If $\operatorname{rank}([\mathbf{G}^\top, \mathbf{1}]) < |\Omega|$, then there exists a length-$|\Omega|$ vector $\boldsymbol{\alpha}$ such that $\boldsymbol{\alpha} \neq 0$

$$\begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix} \operatorname{diag}(\boldsymbol{\mu}(\omega)) \cdot \boldsymbol{\alpha} = 0,$$

where $\operatorname{diag}(\boldsymbol{\mu}(\omega))$ represents the $|\Omega| \times |\Omega|$ matrix with the prior $\mu(\omega)$ on the diagonal. Define matrix

$$\mathbf{Q}' = \mathbf{Q}^* + \mathbf{k} \cdot \boldsymbol{\alpha}^\top$$

with

$$\mathbf{k} = \frac{-\delta \mathbf{Q}^* \boldsymbol{\beta}}{\boldsymbol{\alpha}^\top \cdot \boldsymbol{\mu}'(\omega|s_1^*)}.$$

(Here it is WLOG to assume that the denominator is non-zero. The reason is as follows. WLOG we can assume that $\boldsymbol{\alpha}^\top \cdot \boldsymbol{\mu}^*(\omega|s_1^*) \neq 0$ because if it is equal to zero, we can shift $\boldsymbol{\mu}^*(\omega|s_1^*)$ within the small ball so that it becomes non-zero. Then there must exist small enough δ such that the denominator $\boldsymbol{\alpha}^\top \cdot \boldsymbol{\mu}'(\omega|s_1^*)$ is non-zero.) It is easy to verify that the vector $\mathbf{k}$ we picked satisfies $\mathbf{1}^\top \cdot \mathbf{k} = 0$ because $\mathbf{1}^\top \mathbf{Q}^* = \mathbf{1}^\top$ and $\mathbf{1}^\top \boldsymbol{\beta} = 0$. Therefore by choosing small enough δ, matrix $\mathbf{Q}' = \mathbf{Q}^* + \mathbf{k} \cdot \boldsymbol{\alpha}^\top$ will represent a valid likelihood function $\mu'(s_2|\omega) \in M_2$, because M_2 contains a ball with radius ε centered at $\mathbf{Q}^*$. Let $\mu'(\omega, s_1, s_2)$ be the joint distribution generated by $\mu(\omega)$, $\mu'(s_2|\omega)$ and any $\mu(s_1|\omega) \in M_1$ that will give the desired $\mu'(\omega|s_1^*) = \mu^*(\omega|s_1^*) + \delta \cdot \boldsymbol{\beta}$.

We first show that condition (c) is satisfied. By our construction of $\mathbf{k}$, participant 1's posteriors about participant 2's signal $\mu'(s_2|s_1^*)$ will remain the same as $\mu^*(s_2|s_1^*)$. Because by Lemma D.1,

$$\begin{aligned}
\boldsymbol{\mu}'(s_2|s_1^*) = \mathbf{Q}' \cdot \boldsymbol{\mu}'(\omega|s_1^*) &= (\mathbf{Q}^* + \mathbf{k}\boldsymbol{\alpha}^\top)\boldsymbol{\mu}'(\omega|s_1^*) \\
&= \mathbf{Q}^* \cdot \boldsymbol{\mu}'(\omega|s_1^*) + \mathbf{k}\big(\boldsymbol{\alpha}^\top \boldsymbol{\mu}'(\omega|s_1^*)\big) \\
&= \mathbf{Q}^* \cdot (\boldsymbol{\mu}^*(\omega|s_1^*) + \delta \cdot \boldsymbol{\beta}) - \delta \mathbf{Q}^* \boldsymbol{\beta} \\
&= \mathbf{Q}^* \cdot \boldsymbol{\mu}^*(\omega|s_1^*) \\
&= \boldsymbol{\mu}^*(s_2|s_1^*).
\end{aligned}$$

Notice that this is not equivalent to (c) because we need the posterior about participant 2's report $\mathbf{r}_2$ to be unchanged. However, it suffices to prove that for all $s_2 \in \mathcal{S}_2$, the posteriors $\mu^*(\omega|s_2)$ and

$\mu'(\omega|s_2)$ lead to the same report $\mathbf{G} \cdot \boldsymbol{\mu}^*(\omega|s_2) = \mathbf{G} \cdot \boldsymbol{\mu}'(\omega|s_2)$. By our construction of $\mathbf{Q}'$

$$\mathbf{Q}' = \mathbf{Q}^* + \mathbf{k} \cdot \boldsymbol{\alpha}^\top$$

and our selection of $\boldsymbol{\alpha}$ which guarantees

$$\begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix} \text{diag}(\boldsymbol{\mu}(\omega)) \cdot \boldsymbol{\alpha} = 0,$$

it holds that

$$\begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix} \text{diag}(\boldsymbol{\mu}(\omega)) \cdot (\mathbf{Q}')^\top = \begin{bmatrix} \mathbf{G} \\ \mathbf{1}^\top \end{bmatrix} \text{diag}(\boldsymbol{\mu}(\omega)) \cdot (\mathbf{Q}^*)^\top.$$

The l-th row of the equation imply that for any $s_2 \in \mathcal{S}_2$,

$$\sum_\omega \mathbf{G}[l, \omega](\mu(\omega)\mu'(s_2|\omega)) = \sum_\omega \mathbf{G}[l, \omega](\mu(\omega)\mu^*(s_2|\omega)),$$

where $\mathbf{G}[l, \omega]$ represents the element in row l and column ω of $\mathbf{G}$. This is equivalent to

$$\sum_\omega \mathbf{G}[l, \omega]\mu'(\omega|s_2)\mu'(s_2) = \sum_\omega \mathbf{G}[l, \omega]\mu^*(\omega|s_2)\mu^*(s_2),\ \forall s_2 \in \mathcal{S}_2. \tag{16}$$

The last row means that for all $s_2 \in \mathcal{S}_2$,

$$\sum_\omega \mu(\omega) \cdot \mu'(s_2|\omega) = \sum_\omega \mu(\omega)\mu^*(s_2|\omega),$$

which is equivalent to

$$\mu'(s_2) = \mu^*(s_2),\ \forall s_2 \in \mathcal{S}_2. \tag{17}$$

Combining (16) and (17), we get

$$\sum_\omega \mathbf{G}[l, \omega]\mu'(\omega|s_2) = \sum_\omega \mathbf{G}[l, \omega]\mu^*(\omega|s_2),\ \forall s_2 \in \mathcal{S}_2. \tag{18}$$

This means that $\mu^*(\omega|s_2)$ and $\mu'(\omega|s_2)$ lead to the same report $\mathbf{G} \cdot \boldsymbol{\mu}^*(\omega|s_2) = \mathbf{G} \cdot \boldsymbol{\mu}'(\omega|s_2)$ for any $s_2 \in \mathcal{S}_2$, which completes our proof of condition (c).

Finally, it is easy to show that (a) will also be satisfied. Equations (17) and (18) together imply that the marginal distribution of $\mathbf{r}_2$ remains the same, i.e.,

$$\mu'(\mathbf{r}_2) = \mu^*(\mathbf{r}_2).$$

D.2 Proof of Theorem 5.3

First because of conditional independence we have $\mu(\mathbf{s}_{-i}|s_i) = \sum_{\omega \in \Omega} \mu(\omega|s_i)\mu(\mathbf{s}_{-i}|\omega)$, or equivalently,

$$\mu(\mathbf{s}_{-i}|s_i) = \mathbf{P}_i^\mu \cdot \mu(\omega|s_i).$$

If there exists μ and i such that $\text{rank}(\mathbf{P}_i^\mu) < |\Omega|$, then the columns of $\mathbf{P}_i^\mu$ are linearly dependent. So there exists a non-zero vector $\mathbf{a}$ such that

$$\mathbf{P}_i^\mu \cdot \mathbf{a} = 0.$$

Since the set of possible likelihood function M_i contains a ball with radius ε,[4] there must exists a small enough δ such that $\boldsymbol{\mu}(\omega|s_i) + \delta\mathbf{a}$ is a possible prediction of the state. More specifically, there exists $\widetilde{\mu} \in M$ with the same prior $\widetilde{\mu}(\omega) = \mu(\omega)$ and

$$\widetilde{\mu}(\omega|s_i) = \mu(\omega|s_i) + \delta\mathbf{a}.$$

[4]Here M_i contains a ball means that there is a ball $\mathcal{B} \subseteq \mathbb{R}^{|\mathcal{S}_i| \times |\Omega|}$ such that the intersection of the ball and the space of valid conditional distributions is still in M_i, i.e., $\mathcal{B} \cap (\Delta\mathcal{S}_i)^{|\Omega|} \subseteq M_i$.

In addition, by our definition of conditional independent knowledge, $\widetilde{\mu}$ can further have the same likelihood function for other agents $\widetilde{\mu}(\mathbf{s}_{-i}|\omega) = \mu(\mathbf{s}_{-i}|\omega)$, which leads to

$$\widetilde{\mu}(\mathbf{s}_{-i}|s_i) = \mathbf{P}_i^{\mu} \cdot \widetilde{\mu}(\omega|s_i) = \mathbf{P}_i^{\mu} \left(\mu(\omega|s_i) + \delta\alpha\right) = \mathbf{P}_i^{\mu} \cdot \mu(\omega|s_i) = \mu(\mathbf{s}_{-i}|s_i).$$

This is equivalent to

$$\widetilde{\mu}(\mathbf{r}_{-i}|s_i) = \mu(\mathbf{r}_{-i}|s_i)$$

because $\widetilde{\mu}(\omega|s_j)$ remains the same for all $j \neq i$. So agent i's posterior about others' reports is the same but the truthful report becomes different $\widetilde{\mu}(\omega|s_i) \neq \mu(\omega|s_i)$. This violates the necessary condition for single-tasks problems, i.e., the robust stochastic relevance condition Proposition 3.3, which requires

$$Q_i(\widetilde{\mu}(\omega|s_i)) \cap Q_i(\mu(\omega|s_i)) = \emptyset.$$

It also violates the necessary condition for multi-task problems Proposition 5.3 since we have $\mu(\mathbf{r}_{-i}) = \widetilde{\mu}(\mathbf{r}_{-i})$.